Structural and functional characterization of human ERp44: a closer look at a member of PDI family regulating protein quality control in the early secretory pathway

Thesis

Stefano Vavassori

Structural and functional characterization of human ERp44: a closer look at a member of PDI family regulating protein quality control in the early secretory pathway

Thesis submitted in partial fulfillment of the requirements of the Open University for the degree of

Doctor of Philosophy in Molecular and Cellular Biology

July 2010

Università Vita-Salute San Raffaele
DiBiT
Milan, Italy

DATE of SUBMISSION: 28 APR 2010
DATE of AWARD: 22 JUL 2010

a Mamma e Papà

"La tecnologia progredisce accumulando le esperienze
di molti, non per atti isolati di singoli eroi"

Jared Diamond, "Armi , acciaio e malattie"

"Sono proprio soddisfatto di questo traguardo:
ho appena scoperto il millesimo modo
per non inventare una lampadina !"

Thomas Edison

"Non è difficile dire cose interessanti quando
non si ha l'onere della prova;
il difficile è dire cose interessanti
che siano vere"

Edoardo Boncinelli, "Il cervello, la mente, l'anima"

Declaration

This thesis has been written by myself and has not been used in any previous degree application. All the presented data were obtained by myself, except the experiments in the figures 3.9, 3.11 and 3.12 which were done in collaboration with Wang's group (National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China)

All sources of information are acknowledged by means of reference. Some of the work contained in this Thesis has been published:

Crystal structure of human ERp44 shows a dynamic functional modulation by its carboxy-terminal tail.

Wang L, Wang L, Vavassori S, Li S, Ke H, Anelli T, Degano M, Ronzoni R, Sitia R, Sun F, Wang CC.

EMBO Rep. 2008 Jul;9(7):642-7. Epub 2008 Jun 13.

Table of contents

Abbreviation

2ME 2-mercaptoethanol

CHX cycloheximide

DLS dynamic light scattering

DTT dithiothreitol

DSP (dithiobis)succinimidyl propionate

EDTA ethylenediamine tetracetic acid

ER endoplasmic reticulum

GC golgi complex

ESC early secretory compartment

FPLC fast protein liquid chromatography

Gdn-HCI guanidinium chloride

GFP green fluorescent protein

GST glutatione S-transferase

LB Luria-Bertani

HEPES N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid

HMWC high molecular weight complexes

IAA iodoacetamide

IP_3R inositol trisphosphate receptor

IPTG isopropyl β-D-1-thiogalactopyranoside

NEM N-ethyl maleimide

Ni-NTA nickel-nitrilotriacetic acid

PBS phosphate buffered saline

PCR polymerase chain reaction

PEG polyethylene glycol

SAD single-wavelength anomalous dispersion

SBS substrate binding site

SDM site-directed mutagenisis

SDS sodium dodecyl sulphate

TCA trichloroacetic acid

TCEP tris(2-carboxyethyl) phosphine

TMR thiol-mediated retention

TRIS tris(hydroxymethyl)aminomethane

TRX thioredoxin

UPR unfolded protein response

Table of Figures

Abstract

The Endoplasmic Reticulum (ER) is the site of folding and assembly of secretory proteins. Fidelity of protein-based intracellular communication is guaranteed by protein quality control mechanisms located at the Early Secretory Compartment (ESC), which restricts forward transport to native proteins. ERp44 plays a key role in the Thiol-Mediated Retention (TMR) of variety of client proteins (Ero1, SUMF1, Adiponectin, IgM) thus regulating their transport and localization. Little is known about the molecular mechanisms of ERp44-TMR and how it is regulated in living cells. Hence, the overall aim of this work is to investigate the structure-function relationship of ERp44.

In collaboration with Wang's group the crystal structure of ERp44 was determined. The structure of ERp44 most likely represents a non-reactive conformation of the protein. Indeed, Cys29 is shielded from the bulk solvent by C-terminal tail and almost inaccessible for the formation of intermolecular disulfide bonds with client proteins.

Based on the obtained structural data and functional studies, a panel of mutants of ERp44 has been characterized in order to understand how C-terminal tail rearrangements expose substrate binding site, thus modulating substrates binding/release in view of its role in TMR.

Moreover, the pH gradient between ESC organelles was investigated as major determinant of C-terminal tail rearrangements. Given the known pH differences in the ESC, the data support the hypothesis of a role of the pH variation in governing ERp44-TMR activity *in vivo*.

A deeper knowledge of the structure/function relationship of ERp44 will shed light on the protein quality control mechanisms thus providing essential knowledge of ESC processing diseases and in biotechnology, improving the production of man-made therapeutic proteins.

1. Introduction

1.1 The secretory pathway

All cells secrete a number of molecules to modify their environment in order to protect themselves (or the organism they belong to), to interact with and "communicate" to one another. Smaller molecules are usually transported across the plasma membrane by carriers located in the membrane, whereas almost all secretory proteins in eukaryotes are secreted by exocytosis. In prokaryotes, secretion of proteins is directly mediated by the plasma membrane *via* either signal-recognition-particle-dependent translocation which is homologous to protein translocation into the Endoplasmic Reticulum (ER) in eukaryotic cells (Wolin 1994) or the prokaryote-specific Sec system (Schekman 1994). In contrast, eukaryotic cells have evolved a complex secretory pathway, consisting of several compartments which contain different sets of proteins (ER, Golgi complex, endosomes called endomembrane system) (Becker & Melkonian 1996) (Fig. 1.1). Upon insertion into the ER, proteins travel through this endomembrane system to reach subsequent compartments and then their final destination. The secretory pathway (anterograde pathway) is countered by an endocytotic pathway (retrograde pathway) originating at the plasma membrane. The two pathways are interconnected by crossroads at various steps. Eukaryotic cells, thus, have evolved an integrated intracellular traffic system where vesicular shuttles are believed to be the major transport vehicles.

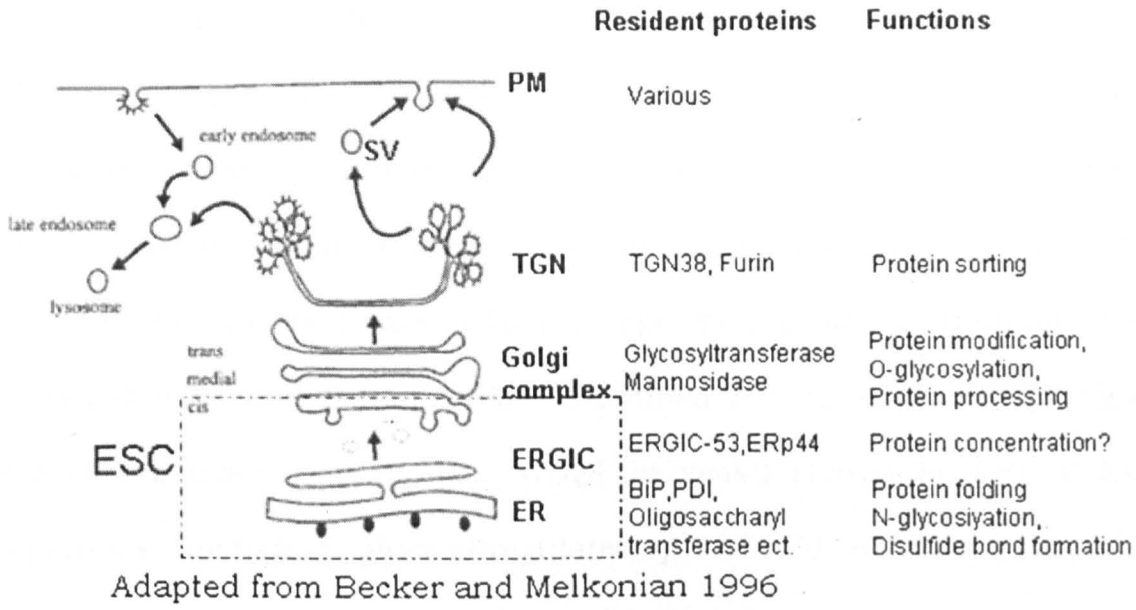


Figure 1.1. Schematic representation of the secretory pathway in eukaryotic cells

TGN: *trans*-Golgi network; SV secretory vesicle; PM: Plasma membrane. Dotted square indicated the ESC. See text for the details

1.1.1 Early secretory compartment (ESC)

1.1.1.1 Endoplasmic reticulum

ER consists of a network of membranous tubules and sacs called cisternae, so extensive that it accounts for more than half the total membrane in many eukaryotic cells. There are two specific regions of ER that are different in structure and function: rough ER (rER) and smooth ER (sER). The former is specialized in protein synthesis while sER has a central role in calcium storage/homeostasis and in the biosynthesis of steroids, cholesterol and other lipids.

The ER, the "gateway" of the secretory pathway, is the site of maturation of secretory proteins. Indeed, through its oxidizing environment and its specialized folding helpers, the ER provides an optimal milieu for protein folding and assembly. In the ER, several co-translational and post-translational modifications take place that do not occur in the cytosol,

such as disulfide-bond formation, N-linked glycosylation and GPI-anchor addition (see 1.2). These covalent changes are important for correct protein folding and assembly (Ellgaard & Helenius 2003). ER resident soluble proteins constitute a special class of proteins that can exit from ER, but eventually are retrieved back by specific receptor. Pelham and co-workers (Munro and Pelham 1987) identified a C-terminal tetrapeptide motif (KDEL motif in mammals, HDEL in yeast) shared by several ER chaperones, including glucose-regulated protein-78 (GRP78), glucose-regulated protein-94 (GRP94) and protein disulphide isomerase (PDI). The KDEL receptor (KDEL-R) is a seven-transmembrane-domain protein that awaits its ligands in the Golgi complex where, once bound the chaperone, sequesters it into vesicle to retrotransport chaperones to the endoplasmic reticulum.

The proof of concept of this process was provided with the addition of the last six amino-acid residues of GRP78 to the C-terminus of an exogenous protein (lysozyme) engineered to be luminal and to cross the secretory pathway; this prevented its secretion, with its consequent accumulation in the ER (Munro and Pelham 1987). Indeed, over-expression of an HDEL-tagged exogenous protein leads to increased secretion of the endogenous yeast chaperone GRP78, indicating that the HDEL retention system can be saturated by HDEL ligands and that chaperone retrieval into the ER is mediated by a specific receptor (Dean & Pelham 1990). Very recently, Sallese and co-workers (Pulvirenti et al., 2008) have reported that KDEL-R triggers a signalling cascade (see ESC signalling pathway 1.5).

1.1.1.2 ERGIC

The ER-Golgi intermediate compartment (ERGIC) is a complex membrane system between the ER and the Golgi complex. Although it is now generally accepted that the ERGIC is an obligatory intermediate compartment in protein transport from the ER to the Golgi complex, the question of whether the ERGIC is only a transient membrane structure (the maturation hypothesis) or a true compartment (the stable compartment hypothesis) remains unanswered (Appenzeller-Herzog & Hauri 2006). According to the maturation hypothesis, the ERGIC clusters would be formed by homotypic fusion of ER-derived, initially COPII coated vesicles, to form a new cis-Golgi cisternae (Bonfanti et al., 1998) (Glick & Nakano 2009). This hypothesis is based mainly on studies at the light microscope level of the anterograde reporter GFP-tagged VSV-G protein in living cells (Presley et al., 1997) (Scales et al., 1997). The results of these studies however also support the stable compartment hypothesis, assuming the ERGIC clusters are highly dynamic. According to the stable compartment hypothesis, the ER-derived vesicles would fuse with ERGIC clusters rather than with themselves. Anterograde and retrograde traffic from the ERGIC would not entirely consume the clusters, but part of them would remain and function as an acceptor for new membrane traffic from both the ER and the *cis*Golgi.

On the whole, in the most popular current view, major role of the ERGIC is the sorting of anterograde and retrograde traffic (Klumperman et al. 1998) (Bannykh et al., 1998) (Otte & Barlowe 2004) (Appenzeller-Herzog & Hauri 2006) (Glick & Nakano 2009).

1.1.1.3 Golgi complex

In mammalian cells, Golgi Complex (GC) consists of flattened membrane-bound compartments, called cisternae, which form Golgi stacks, themselves interconnected by lateral tubules to form the Golgi ribbon. Both Golgi stacks and ribbon are polarised with an entry (*cis*) face, where cargo molecules synthesised in the ER reach the Golgi complex, and an exit (*trans*) face, where they leave for downstream localizations (Pelham & Rothman 2000).

The GC is situated at the "heart" of the secretory pathway, and its main functions are to modify (mainly through glycosylation) and sort proteins. Glycan addition can be again subdivided into two main groups: N- and O-glycosylation. The former is the addition of sugars to the amide nitrogen of asparagine side chains, while the latter to the hydroxyl groups of serines or threonines. N-glycosylation of proteins starts at the cytosolic face of the ER by the synthesis of a GlcNAc and mannose (Man) containing glycan chain onto the lipid dolichol phosphate. This chain is then flipped into the lumen of the ER, where it undergoes a further addiction of Man and glucose (Gluc) residues. Eventually, oligosaccharyltransferase moves GlcNAc2Man9Gluc3 moieties onto suitable asparagine side chains of proteins co-translationally (see also 1.2). Elaboration of these glycan chains is performed in the GC to create an array of different structures classified into three categories. High mannose type chains end in mannose residues just like the ER exported form, and are similar to those prevalent in lower eukaryotes such as yeast. These mainly exist on a few specialized mammalian proteins. In contrast, to generate complex type N-glycans

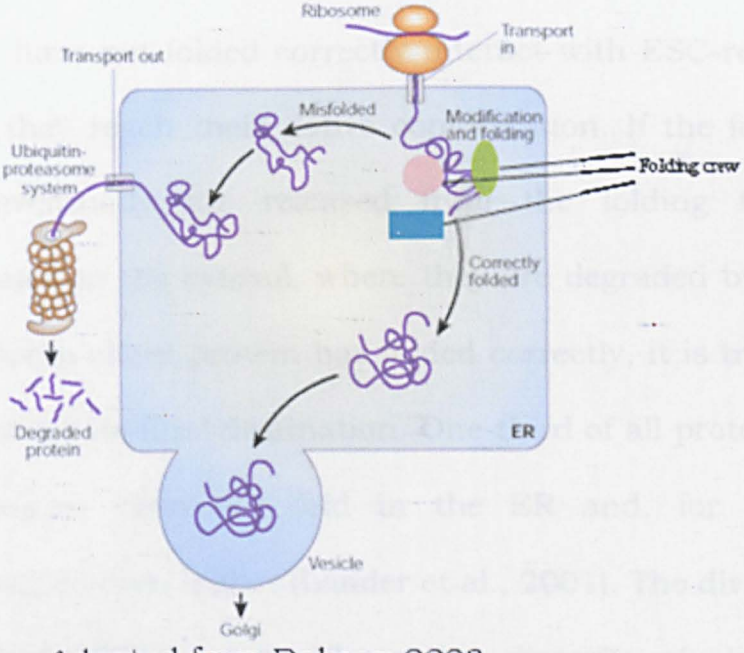
most of the ER derived mannose residues are trimmed in the early Golgi. Addition of GlcNAc residues then generates further branches on the chain. To this galactose (Gal), sialic acid (NeuAc) and/or fucose (Fuc) are added in the late Golgi to create the final complex glycan product. The exact mechanism of O-glycosylation is less well established. The main forms of Golgi-derived O-glycans on higher eukaryotic secretory proteins are mucin type glycans and proteoglycans.

Many glycosylation enzymes are expressed cell type specifically, restricting the choice to a subset of glycosyltransferase reactions in any given cell (Ungar 2009). Glycosyltransferases and glycosidases are sub compartmentalized; different cisternae of the Golgi complex contain a different set of enzymes (Ungar 2009). This allows enzymatic reactions to be operated sequentially, thus secretory proteins move along the Golgi complex like on a conveyor belt while reaching their final form.

The endoplasmic reticulum, ERGIC and *cis*Golgi complex, called collectively Early Secretory Compartment (ESC)(Fig.1.1), have a stringent quality-control system for "proof-reading" newly synthesized proteins, so that only native conformers can leave the ESC reaching their final destinations (Anelli & Sitia 2008). Non-native conformers and incompletely assembled oligomers are retained, and, if misfolded persistently, they are degraded (Ellgaard & Helenius 2003) (Braakman & Sitia 2003). The ESC support the folding, assembly, post-translational modification and quality control as well just like a molecular "protein factory".

1.2 ESC protein folding crew

Protein folding of a newly synthesized protein can start as soon as the N-terminus of the nascent peptide emerges from the translocon Sec61. A protein may be able to reach its native conformation without assistance, but this is unlikely in the crowded environment of the cell where the risk of aggregation is high. Therefore, a "folding crew" is present to take care of newly synthesized proteins (Dobson 2003) (Fig. 1.2).



Adapted from Dobson 2003

Figure 1.2. Regulation of protein folding in the ER.

Secretory proteins are translocated into the ER, where they gain their native structure with the help of ESC folding crew. Properly folded proteins are then transported forward. Non-properly folded proteins are detected by a quality-control mechanism and degraded by cytosolic proteasomes.

This crew can catalyze slow folding steps, prevent unproductive interactions with other proteins, or prevent proteins from getting trapped in off-pathway intermediates. A general danger during protein folding, whether in the cytosol, ESC or mitochondria, is the exposure of hydrophobic residues, which may form undesirable intra-chain or inter-

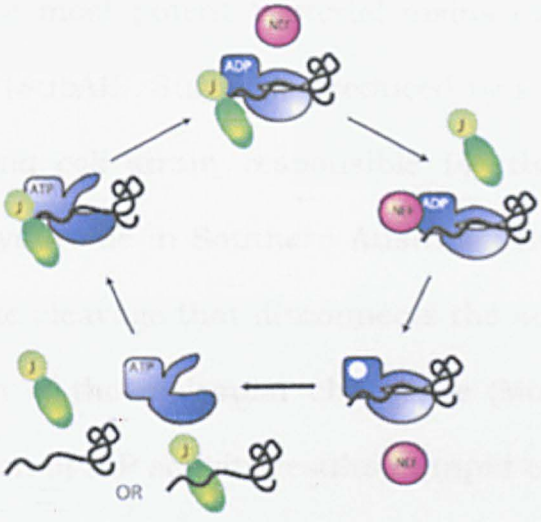
chain interactions, leading to misfolding and often aggregation. Hsp70 (-like) chaperones present in all cellular compartments help to prevent this, keeping newly synthesized proteins in a folding competent state (Bukau et al., 2006). Protein folding in the ER involves two additional features that distinguish the process from folding in the cytosol, namely the formation of disulfide bonds, and attachment of N-linked glycans to polypeptides. Specialized chaperones and folding enzymes are involved in these processes (see also 1.2.2 and 1.2.3).

Proteins that have not folded correctly interact with ESC-resident folding factors until they reach their native conformation. If the folding process fails, they eventually are released from the folding factors to be retrotranslocated to the cytosol, where they are degraded by proteasomes (Fig. 1.2). When a client protein has folded correctly, it is transported out of the ER towards its final destination. One-third of all proteins expressed in *Saccharomyces cerevisiae* fold in the ER and, for humans, this percentage may be even higher (Lander et al., 2001). The diverse repertoire of ESC-resident folding crew reflects this diversity of clients: multiple members have been identified for several families of chaperones and folding enzymes (Fig. 1.11).

1.2.1 The classical chaperones

Classical chaperones are grouped in several subfamilies HSPs of 40, 60, 70, 90, and 100 kDa in size. The ER possesses a member of the Hsp70 family (glucose-regulated protein (GRP)78/BiP (Haas IG and Wabl M.1983, (Brewer & Hendershot 2005), a chaperone of the hsp70 family that plays a key regulatory role in ER signalling UPR (see 1.4). BiP was first isolated as

a protein associating with unassembled Ig-H chains (Haas and Wabl, 1983). Peptide binding studies have confirmed that BiP has a preference for peptides with aliphatic residues, which are usually buried in properly folded/assembled proteins (Blond-Engluidi et al., 1993) (Foresti et al., 2003). Like other Hsp70s, BiP has an N-terminal ATPase domain and a C-terminal substrate binding domain. These domains communicate, as cycles of ATP hydrolysis and ADP-to-ATP exchange are coupled to cycles of substrate binding and release (Mayer et al., 2003) (Christis et al., 2008) (Fig. 1.3). The interdomain linker is crucial in communicating substrate and nucleotide binding from one domain to the other, an event that is accompanied by major conformational changes in both domains (Jiang et al., 2005) (Awad et al., 2008). Owing to the weak BiP ATPase activity, Hsp40-like co-chaperones containing J domains (ERdj) play a key regulatory role. Five ERdj proteins have been isolated so far (Cunnea 2002) (Kroczynska et al., 2005) (Shen & Hendershot 2005). One of them, ERdj5, also displays oxidoreductase activity, possibly linking BiP-dependent folding/unfolding and disulfide bond formation, isomerisation or reduction (Ushioda et al., 2008).



From Christis et al., 2008

Figure 1.3. The chaperone BiP ATPase cycle.

The cycle starts by the binding of substrate, which may be presented by one of the five J proteins in the ER. J then stimulates BiP's ATPase activity and bound ATP is hydrolyzed, leading to a conformational change in BiP, which closes the lid domain and drastically decreases the on and off rates of substrate from BiP. One of the two nucleotide exchange factors then mediates the release of ADP, allowing the binding of ATP, which opens the lid to release the substrate for another round.

BiP levels are tightly controlled and are elevated upon physiological (e.g., cell differentiation increasing the ER load with cargo proteins) or pathological (accumulation of misfolded proteins) stress responses.

During embryogenesis, BiP expression remains below detection limit to the morula stage, but becomes abundant at the blastocyst stage in E3.5 embryos (Kim et al., 1990). It is therefore not surprising that the *BiP*$^{-/-}$ mice cannot survive beyond this stage, and the cells when cultured in vitro cannot proliferate and rapidly degenerate (Luo et al., 2006). A partial reduction in the level of BiP is tolerated as *BiP*$^{+/-}$ embryos and adult mice are normal. The essential role of BiP in cell, tissue, and organism homeostasis is underscored by the identification of BiP as the cellular

target of one of the most potent bacterial toxins ever characterized the subtilase cytotoxin (SubAB). SubAB is produced by a highly virulent Shiga toxigenic *Escherichia coli* strain responsible for the 1998 outbreak of hemolytic uremic syndrome in Southern Australia. SubAB inactivates BiP through a single-site cleavage that disconnects the substrate binding from the ATPase domain of the molecular chaperone (Montecucco & Molinari 2006). The disruption of BiP activity results in rapid cell death.

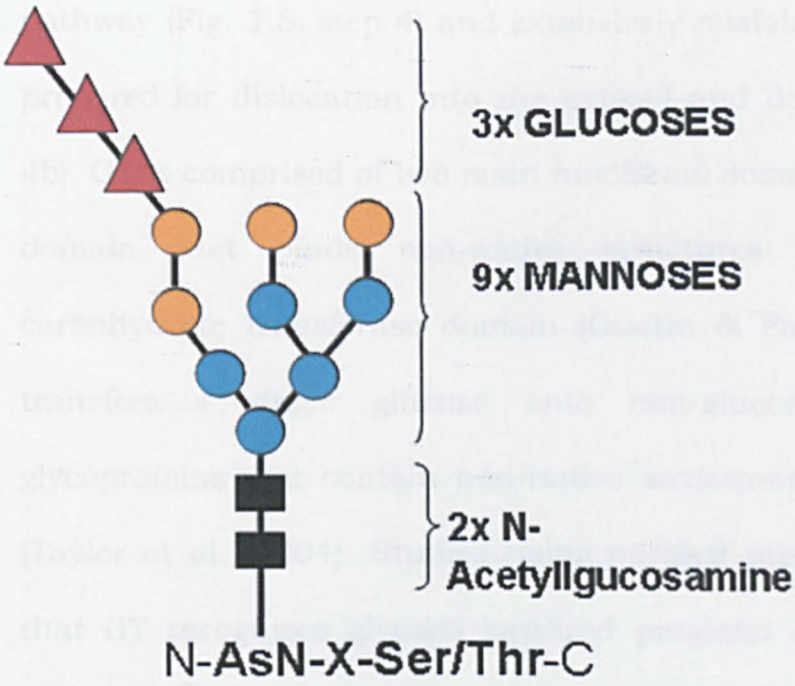
1.2.2 *The lectin chaperones*

The majority of secreted proteins are modified by multiple *N*-linked glycans. N-glycans play key roles: during folding they dictate the association with lectin chaperones, increase the solubility of the polypeptide. In contrast to BiP that binds directly to the hydrophobic backbone of the polypeptide, the lectin chaperones bind the glycans. Once the protein is folded, glycans participate in many key biological processes, such as self/non-self recognition in immunity, signal transduction, and cell adhesion (Ohtsubo & Marth 2006).

N-linked glycosylation of asparagine residues in an N–X–S/T motif is an ER-specific protein modification. Preformed oligosaccharide units (GlcNAc2-Man9-Glc3) are co-translationally transferred *en bloc* to the nascent protein by the oligosaccharyltransferase complex (Hebert & Molinari 2007) (Fig. 1.4). Indeed, as folding proceeds, glycan acceptor sites can become buried and remain unmodified. This suggests folding and glycosylation compete with each other *in vivo* (Holst et al., 1996) (Foresti et al., 2003).

Glucose trimming by glucosidases I and II produces a monoglucosylated species that can bind to the lectin chaperones calnexin and calreticulin

(Hammond & Helenius 1994). The two proteins are highly homologous, apart from the fact that calnexin is a transmembrane protein and calreticulin is soluble. Calnexin is thought to interact with glycans closer to the membrane, whereas calreticulin binds more peripheral glycans (Molinari 2007). Although both proteins could potentially associate with soluble and membrane proteins, they indeed, interact with a distinct set of client proteins. This may partly be the result of their different localization in the ER because, when the transmembrane segment of calnexin was fused to calreticulin, the pattern of associating proteins shifted towards that normally seen for calnexin (Molinari & Helenius 1999) (Molinari 2007).



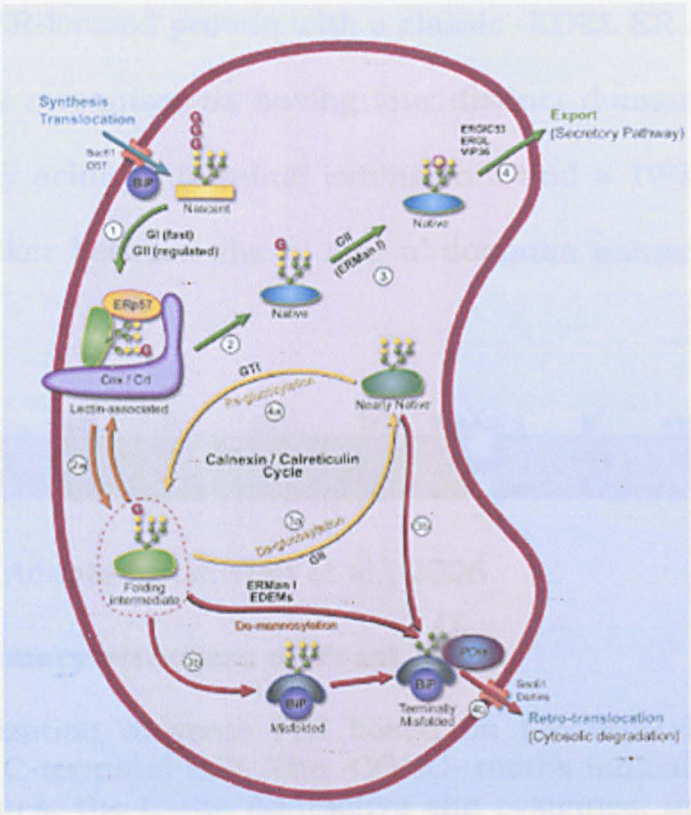
Adapted from Hebert and Molinari 2007

Figure 1.4. Structure of *N*-linked oligosaccharides.

The high mannose precursor covalently attached to Asn-X-Ser/Thr sequences of nascent polypeptide chains normally contains 9 mannose residues and 3 glucoses.3 Mannose (orange circle) and 3 glucoses are further removed during processing and quality control. See text for the details.

Upon substrate release from calnexin (Fig. 1.5, steps 2 and 2a), glucosidase II also removes the final glucose (glucose 3) creating the unglucosylated substrate, thus inhibiting substrate rebinding to the lectin chaperones (Fig. 1.5, step 3 for native proteins, step 3a for folding intermediates, and step 3b for folding-incompetent polypeptides). Regeneration of the monoglucosylated state and rebinding to the lectin chaperones is controlled by the UDP-glucose: glycoprotein glycosyltransferase (GT; acronyms such as UGT or UGGT are also commonly used), and it is only possible for nearly native folding intermediates (Fig. 1.5, step 4a) (Caramelo et al., 2003) (Caramelo et al., 2004). GT ignores native proteins that are released into the secretory pathway (Fig. 1.5, step 4) and extensively misfolded polypeptides that are prepared for dislocation into the cytosol and degradation (Fig. 1.5, step 4b). GT is comprised of two main functional domains: a large NH_2-terminal domain that binds non-native structures and a COOH-terminal carbohydrate transferase domain (Guerin & Parodi 2003). This protein transfers a single glucose onto non-glucosylated side chains of glycoproteins that contain non-native structures (Caramelo et al., 2004) (Taylor et al., 2004). Studies using purified proteins have demonstrated that GT recognizes glycans localized proximal to the misfolded domain (Ritter & Helenius 2000). However, another study found that GT could modify a glycan that was separated from the protein defect by at least 40 Å (Taylor et al., 2004). GT provides the essential connection that links the non-native exposed hydrophobic properties of a maturing protein to the composition of its exposed hydrophilic modification responsible for recruiting chaperones. Proteins do not cycle between UGGT and

calnexin/calreticulin indefinitely, however, and those who fail to fold need to be removed from the ER (Hebert & Molinari 2007) (Fig.1.5).



From Hebert and Molinari 2007

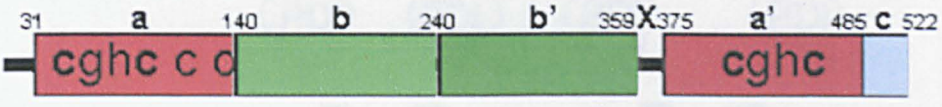
Figure 1.5. The fate of newly synthesized glycoproteins in the ER lumen. See text for details.

1.2.3 *Oxidative folding*

The ER lumen contains various oxidoreductases assisting in disulfide bond formation in newly synthesized proteins. Disulfide bonds play a role in protein folding/assembly, in the stabilization of secretory proteins and their folding intermediates (structural disulfide bonds) but, also in the regulation of the function of certain proteins (functional disulfide bond) (Danon 2002).

1.2.3.1 Protein disulfide isomerase (PDI)

PDI was the first folding catalyst ever reported (Venetianer and Straub 1963). It is an ER-located protein with a classic -KDEL ER retrieval motif. PDI is currently recognized as having four distinct domains, a, b, b', and a', plus a highly acidic C-terminal extension c and a 19-amino-acid long interdomain linker between the b' and a' domains named x (Tian et al., 2006) (Fig. 1.6).



Adapted from Tian et al., 2006

Figure 1.6. Primary structure of Yeast PDI

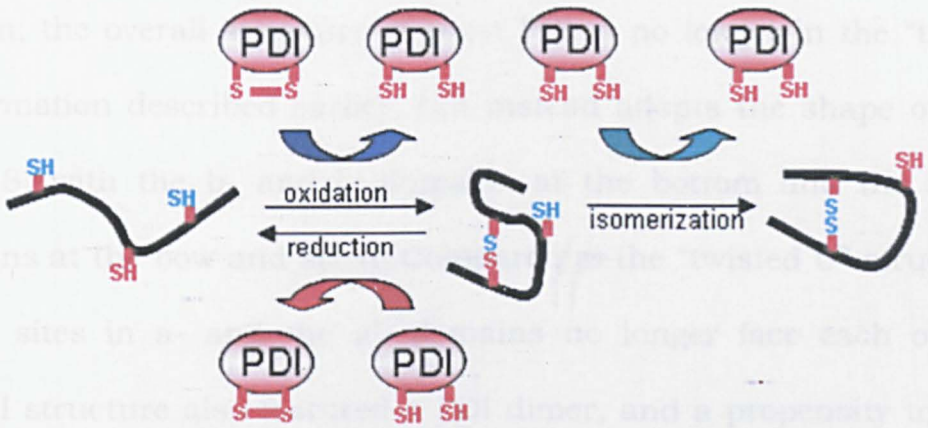
Domain organization of yeast PDI based on the crystal structure (c represents the C-terminal tail). The -CGHC- motifs indicate the location of the active sites, the C the non-active site cysteines, and X the loop connecting the b' and a' domains.

The active sites of most PDI family members consist of a -CGHC- motif (Cys, Gly, His, Cys). When PDI is in the oxidized state, the disulfide can be transferred to the substrate to catalyze its oxidation, whereby the active site itself becomes reduced. When in the reduced state, substrate disulfides can be reduced and the active site ends up in the oxidized state (Hebert & Molinari 2007) (Fig. 1.7). These thiol–disulfide exchange reactions proceed through the formation of a transient mixed disulfide between enzyme and substrate.

Since reduction requires the second cysteine to resolve the mixed disulfide with the substrate, mutation of this residue leads to the accumulation of

covalent complexes through an intermolecular disulfide (Walker et al., 1996).

The PDI also catalyze the rearrangement of incorrectly-formed disulfides in the process called isomerization (Fig.1.7). Since non-native disulfides are often formed during folding and prevent the formation of the native structure (Jansens et al., 2002), isomerization is a crucial reaction.



Adapted from Hebert and Molinari 2006

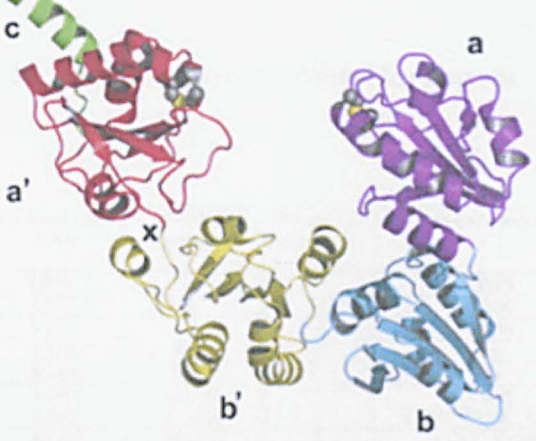
Figure 1.7. PDI oxidation, reduction and isomerization of disulfide bonds. See text for details.

In addition to its role in disulfide bond formation, PDI is the β subunit of prolyl-4-hydroxylase (Koivu et al., 1987) and microsomal triglyceride transfer protein (Wetterau et al., 1991). PDI is also implicated in peptide loading onto MHC class I (Peaper & Cresswell 2008), and in regulating NAD(P)H oxidase (Janiszewski et al.,2005) in the ER. Recently, a number of reports indicated that PDI is, indeed, a marker for the release of intracellular contents from damaged cells (for example, activating tissue factor via catalysis of thiol-disulfide exchange and thus initiating the blood-clotting cascade at the site of wound damage) (Reinhardt et al. 2008).

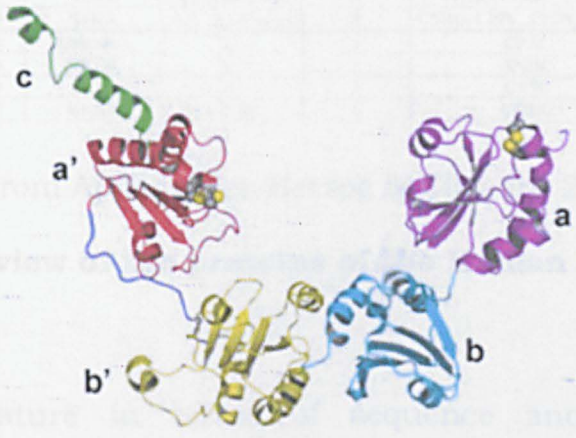
The crystal structure of yeast PDI (Tian et al., 2006) revealed that the four thioredoxin domains are arranged in the shape of a 'twisted U', with the two active sites (in domain a and a'-) facing each other from a distance of 28 Å (Fig. 1.8). A highly hydrophobic patch in the b' domain together with hydrophobic areas surrounding the active sites, form a continuous hydrophobic surface across the a-, b'-, and a' domains (Tian et al., 2006). Recently the same group describes an alternative conformation of the protein; the overall structure of yeast PDI is no longer in the "twisted U" conformation described earlier, but instead adopts the shape of a "boat" (Fig.1.8) with the b- and b' domains at the bottom and the a- and a' domains at the bow and stern. Compared to the "twisted U" structure, the active sites in a- and the a'- domains no longer face each other. The crystal structure also featured a PDI dimer, and a propensity to dimerize in solution and in the ER was confirmed by cross-linking experiments and the split green fluorescent protein system (Tian et al., 2008). In the dimer observed in the "boat" structure, both the active site in a' domain and the potential Substrate Binding Site (SBS) in the b' domain are buried. This conformation therefore presumably represents an "inactive" state of PDI, because most of its catalytic elements are inaccessible (Tian et al., 2008). On the other hand, the "twisted U" conformation most likely represents an "active" form of PDI since its catalytic elements are exposed to the solvent. The transition between monomeric and dimeric PDI could be a regulatory mechanism whereby the ER adjusts its folding capacity to meet different cellular demands.

Another conformational change reported in PDI (human PDI- Nguyen et al., 2008) is linked to substrate-binding site availability in domain b'. In

both yeast PDI and human PDI, the b' domain is thought to be a primary substrate binding site (Klappa et al., 1998) (Tian et al., 2006). Biophysical and NMR analysis indicated that the x region could move from a conformer in which it capped the SBS in domain b', to that observed in the crystal structure of yeast PDI, in which the SBS was uncapped (Nguyen et al., 2008). Such change in conformation could be linked to the substrate-binding cycle and may also modulate the specificity of substrate binding by PDI.



From Tian et at., 2006



From Tian et at., 2008

Figure 1.8. Overall structures of yPDI.

Ribbon diagram of PDI with the a, b, b', and a' domains in magenta, cyan, yellow, and red, respectively, and the C-terminal extension in green. The side chains of the active site cysteines in the a and a' domains are showed in spacefilling representation with the sulfur atoms in yellow. The top picture is "twisted U" structure (accession number 2B5E) while below is "boat" structure (accession number 3B0A).

1.2.3.2 PDI Family

Although PDI was for a long time considered as the sole enzymatic catalyst of thiol-disulfide exchange reactions in the ESC, it is now clear that other PDI homologues are important in this process. Indeed, 19 other ESC-resident proteins with at least one thioredoxin-like domain have been

identified so far in humans (Appenzeller-Herzog & Ellgaard 2008) (Fig. 1.9).

Name	Accession	Length	ER-localization motif	Domain composition	Number of a-type domains	Active-site sequence	PDB accession numbers[c]
Hag 3	Q8TD06	165	QSEL[b]	a	1	CQYS	
ERp18	O95881	172	EDEL	a	1	CGHC	1sen
Hag 2	O95994	175	KTEL[b]	a	1	CPHS	
ERp28/29[d]	P30040	261	KEEL	b–D	0	n.a.	1g7e, 2e0e, 1ovn
ERp27	Q96DN0	273	KVEL	b–b′	0	n.a.	
ERp44	Q9BS26	406	RDEL	a–b–b′	1	CRFS	
ERp46[d]	Q8NBS9	432	KDEL	a°–a–a′	3	CGHC, CGHC, CGHC	a°: 2diz
P5	Q15084	440	KDEL	a°–a–b	2	CGHC, CGHC	a: 2dml; a°: 1x5d
ERp57	P30101	505	QEDL	a–b–b′–a′	2	CGHC, CGHC	a: 2alb; b–b′: 2b8l; a′: 2dmm
PDI	P07237	508	KDEL	a–b–b′–a′	2	CGHC, CGHC	a: 1mek; b: 2bjx; a′: 1x5c; Pdi1p: 2b5e
PDIr	Q14554	519	KEEL	b–a°–a–a′	3	CSMC, CGHC, CPHC	
PDIp	Q13087	525	KEEL	a–b–b′–a′	2	CGHC, CTHC	
PDILT	Q8NX07	584	KEEL	a–b–b′–a′	2	SKQS, SKKC	
ERp72	P13667	645	KEEL	a°–a–b–b′–a′	3	CGHC, CGHC, CGHC	a: 2dj1; a′: 2dj2; a°: 2dj3
ERdj5	Q8NXB1	793	KDEL	J–a°′–b–a′–a–a′	4	CSHC, CPPC, CHPC, CGPC	
TMX	Q9H3N1	280	Unknown	a	1	CPAC	1x5c
TMX2[e]	Q9Y320	296	KKDK	a	1	SNDC	2dj0
TMX4	Q9H1E5	349	RQR	a	1	CPSC	
TMX3	FLJ20793	454	KKKD	a–b–b′	1	CGHC	

From Appenzeller-Herzog & Ellgaard 2008

Figure 1.9. Overview of the proteins of the Human PDI family

The unifying feature in terms of sequence and structure is the thioredoxin-like domains that can be either catalytic or non-catalytic. These two types of domains are called a- and b-type domains, respectively, and are distinguished based on sequence similarity. The a-type domains usually contain two cysteines in a CXXC active-site motif, with an intervening GH sequence being the most common in the PDIs. The b-type domains do not have cysteines at the active site and are therefore not redox active. Most human PDIs are soluble proteins with a C-terminal KDEL-like ER-localization motif. Four PDIs are transmembrane proteins of the TMX (for "thioredoxin-related transmembrane protein") subgroup. TMX2 and TMX3 have typical C-terminal KKXX ER localization motifs. The family members differ from PDI in domain organization, tissue specificity and / or sequence of the active site (Hatahet et al.,2009). The multitude of PDI family members reflects both the importance and difficulty of

introducing correct disulfide bonds into client proteins. All of the different family members may have their own expertise in assisting either specific clients or different stages in the folding process. Indeed, Winther and co-workers (Nørgaard et al., 2001) have showed that, in *S. cerevisiae*, the five PDI homologs are not functionally interchangeable. In mammalian cells, the differences between PDI family members are illustrated by the opposing roles played by PDI and ERp72 in retro-translocation. Forster et al. (Forster et al., 2006) found that PDI facilitated retro-translocation of cholera toxin and misfolded protein substrates, whereas ERp72 mediated their retention in the ER.

ERp57 (also known as ERp60, ERp61, GRP57, GRP58, HIP-70, and Q-2) is the closest homolog of PDI.

It is very similar in length and domain architecture to PDI, with the most striking difference being that the c region of PDI is highly acidic, whereas ERp57 contains multiple lysine residues. As a result of its interaction with the two related, ER lectin chaperones CNX and CRT (Ellgaard & Frickel 2003). ERp57 acts on glycosylated substrates (Oliver 1997) (Molinari & Helenius 1999). ERp57 is a multifunctional oxidoreductase that readily catalyzes dithiol oxidation, disulfide isomerization, and reduction in vitro (Frickel et al., 2004) (Lappi et al., 2004). ERp57 has been most widely studied for its role in the immune system because of its involvement in the assembly of the heavy chain of major histocompatability complex (MHC) class I molecules (Wearsch & Cresswell 2008) (Dong et al., 2009). However, it also has a role in ESC quality control for other newly synthesized glycoproteins and has been implicated in the enzymatic

functions of phospholipase C, carnitine palmitoyl transferase, oxidase and reductase (Coe & Michalak 2010). By a combination of a cross-comparison with PDI a binding site in the b' domain of ERp57 was identified as the lectin P-domain interaction site (Russell et al., 2004). This result implied that the b' domains of PDI and ERp57 have become specialized in function relating to direct and indirect substrate recognition.

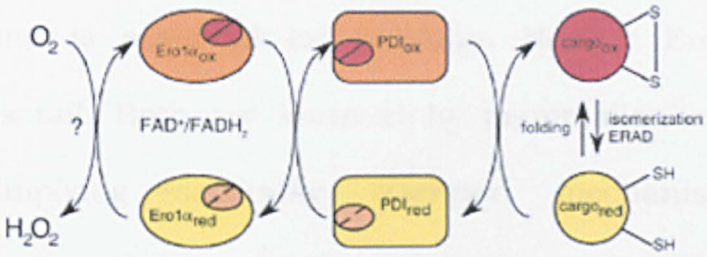
1.2.3.3 Endoplasmic reticulum oxidoreductin 1 (Ero1)

The active site of PDI needs to be recharged after oxidizing a client protein. A long-standing debate on how this is accomplished was terminated by the identification of Ero1p (ER oxidoreductin is a glycosylated flavoenzyme) in a screen for yeast mutants defective in disulfide bond formation (Frand & Kaiser 1998) (Pollard et al., 1998).

The characteristic elements of both yeast and mammalian Ero1 proteins are the bound flavin cofactor FAD, a catalytic CXXCXXC motif and a thioredoxin-like dicysteine motif.

Ero1 and PDI form the major pathway for protein disulfide bond formation in the eukaryotic ER (Masciarelli & Sitia 2008) (Fig. 1.10). Ero1 influences protein oxidation by coupling the oxidizing power of molecular oxygen and its flavin cofactor to generate disulfide bonds (Tu & Weissman 2004) (Gross et al., 2006). These disulfides are transferred from Ero1 to the soluble disulfide-carrier PDI, which directly transfers its disulfides to folding secretory proteins (Frand & Kaiser 1999) (Sitia et al., 2000) (Sitia et al., 2001). Transfer of electrons from substrate proteins to PDI and from PDI to Ero1, proceeds as a series of direct thiol-disulfide exchange reactions (Sevier & Kaiser 2008). Recently, Wang and co-workers

reconstituted in vitro reconstitution of the Ero1-PDI-substrate oxidation system (Wang et al., 2009).



Adapted from Masciarelli and Sitia 2008

Figure 1.10. Ero oxidative folding.

In eukaryotic cells, oxidative folding of cargo proteins is based on serial disulfide interchange reactions involving PDI and Ero1 molecules. *In vitro,* yeast Ero1 is oxidized by molecular oxygen generating H_2O_2 in stoichiometric amounts to the disulphide bonds formed.

This pathway is fundamentally conserved in higher eukaryotes, although some important differences exist between yeast and mammalian cells. First, whilst a single Ero1 gene exists in *Saccharomyces cerevisiae*, two isoforms, Ero1α and Ero1β, are found in mammals. Both are able to complement yeast Ero1 defective mutants, underscoring the conservation of the oxidative folding pathways. Ero1α is induced by hypoxia (Gess et al. 2003) and PPARγ and inhibited by SIRT1 (Qiang et al., 2007). In contrast, Ero1β increases during the UPR (Pagani et al., 2000), and is constitutively expressed at high levels in certain secretory tissues (Dias-Gunasekara et al., 2005), confirming that ER stress responses are important for the acquisition of the secretory phenotype (Brewer & Hendershot 2005) (Masciarelli & Sitia 2008). The reasons underlying the functional differentiation between Ero1α and Ero1β in mammals are not yet clear.

Another remarkable difference between yeast and mammalian Ero1 molecules concerns the mechanisms of sub-cellular localisation. A C-terminal tail is present in yeast Ero1p that mediates membrane association and is essential for function. Neither Ero1α nor Ero1β possesses this tail. Both are secreted by mammalian cells when over-expressed, implying saturable retention mechanism(s). It was demonstrated that the localisation of Ero1α and Ero1β depends on dynamic interactions with PDI or ERp44 (Otsu et al., 2006), two soluble proteins that despite the presence of ER localisation motifs (KDEL and RDEL, respectively) are differentially distributed in the early secretory pathway (Masciarelli and Sitia 2008).

1.3 ESC Quality control

In the late 1980s, work on oligomeric viral proteins (Kreis and Lodish, 1986)(Gething and Sambrook, 1989), the T-cell receptor (Bonifacino et al., 1989)(Sanchoet et al., 1989) and immunoglobulins (Bole et al., 1986)(Sitia et al., 1987) (Hendershot and Kearney, 1988) revealed that proper folding and assembly are requisites for transport to the Golgi complex and onwards along the secretory route. Klausner referred to this phenomenon as 'Architectural editing'. The term 'ER Quality Control' (ER QC) (Hurtley and Helenius 1989) eventually has been proposed to indicate the processes of conformation-dependent molecular sorting of secretory proteins. The aberrant proteins are eventually retro-translocated (or dislocated) across the ER membrane for degradation by cytosolic proteasomes (ERAD). The identity of the retro-translocation channel(s)/dislocon(s) remains controversial as a number of ER

membrane proteins have been proposed to serve this function for the ERAD pathway including Sec61, Derlin and E3 ubiquitin ligase family members (Nakatsukasa & Brodsky 2008). The term "ESC quality control" has been becoming effective since emerging evidence indicated that also ERGIC and *cis*Golgi are involved in protein quality control (Anelli et al. 2007) (Anelli & Sitia 2008).

ESC protein QC is intimately linked to the processes of folding/assembly (Ellgaard and Helenius, 2003) (Sitia and Braakman 2003) (Fig. 1.11). Both rely on chaperones and devoted resident enzymes. QC serves different roles (Anelli and Sitia 2008):

(i) It prevents the deployment of aberrant protein conformers, ensuring that only native proteins proceed along the secretory pathway;

(ii) it retains precursors in an environment suitable for their maturation;

(iii) it increases their local concentration to favour assembly and polymerization;

(iv) it reduces the risks of proteotoxicity by inhibiting aggregation and degrading terminally misfolded proteins;

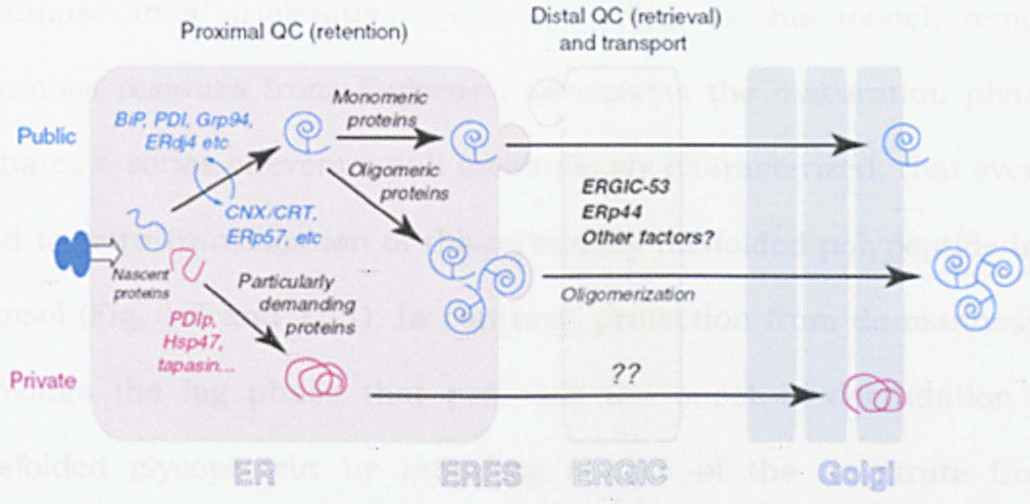
(v) it maintains homeostasis in the early secretory pathway;

(vi) it is involved in the developmental regulation of protein secretion (IgM, adiponectin, see below)

(vii) it is important for storing proteins for regulated secretion. In plants, in fact, ER retention/accumulation is utilized to store abundant proteins during seed formation (Mainieri et al., 2004) (Jolliffe et al., 2005).

A crucial part of the ESC quality control is therefore the recognition of terminally misfolded chains, their extraction from cycles of futile folding

attempts, their unfolding and transport into the cytosol for proteasome mediated degradation (Meusser et al., 2005).



From Anelli and Sitia 2008

Figure 1.11. ESC Quality control

Molecular chaperones help secretory proteins to undergo their native structure. Certain proteins that are produced in large amounts, or are intrinsically difficult to fold, are assisted by specific (private) chaperones and enzymes. QC can occur in sequential steps. Indeed, after a **proximal** QC, certain substrates (generally multimeric proteins) seem to undergo also **distal** QC checkpoints in ERGIC and *cis*-Golgi. This model could mediate cargo concentration and selective export of oligomerized species, thus coupling fidelity and efficiency in the secretory protein factory. While proximal QC can rely on simple retention, the distal checkpoints likely imply substrate retrieval to the ER, either for further attempts to fold, or for retro-translocation and degradation.

Protein quality control system is coordinated by a number of molecular chaperones that monitor non-native determinant in the newly synthesized proteins such as:

1. Presence of immature glycans;

2. Exposure of hydrophobic patches;

3. Exposure of reactive thiols (Thiol-mediated retention).

The finding that inhibition of mannose removal from *N*-glycans protects folding-defective polypeptides from ERAD led to the concept of the 'mannose-timer' (Helenius, 1994). According to this model, removal of mannose residues from *N*-glycans terminates the maturation phase and initiates a series of events, still incompletely characterized, that eventually lead to retro-translocation of the terminally misfolded polypeptide into the cytosol (Fig. 1.5 and 1.11). In contrast, protection from de-mannosylation prolongs the lag phase that precedes the onset of degradation of the misfolded glycoprotein by retarding release of the substrate from the calnexin chaperone system (de Virgilio et al., 1999) (Molinari et al., 2002).

Numerous studies with inhibitors in mammalian cells revealed the active involvement of kifunensine-sensitive 1,2- mannosidase(s) in ERAD regulation (Cabral et al., 2001) (Helenius & Aebi 2004) (Hebert et al., 2005) (Lederkremer 2009).

Removal of this mannose residue, the only acceptor for UGT-catalyzed protein re-glucosylation, thus seems to be a potent signal for glycoprotein disposal in mammalian cells because it results in the irreversible interruption of futile folding attempts in the calnexin cycle. Which protein(s) perform the extensive mannose trimming observed during preparation of the folding defective polypeptides for proteasomal disposal is still a matter of debate and extensive research. At least three possibilities exist:

(1) co-sequestration of ERAD candidates and ER mannosidase I in specialized sub-regions of the ER (Frenkel et al., 2003) in which the mannosidase concentration reaches levels similar to those showed in vitro to cause extensive mannose removal (Herscovics, et al., 2002)

(2) intervention of Golgi endo-mannosidase(s);

(3) intervention of a new class of recently characterized mannosidase-like proteins, EDEM1, EDEM2 and EDEM3 (Molinari et al., 2003) (Oda et al., 2003) (Hirao et al., 2006) (Mast & Moremen 2006).

The recognition of non-glycosylated ERAD substrates has received less attention, but recently two studies have showed that, as non-glycoproteins are substrates of GRP94 or BiP, their ERAD pathways do not completely overlap with those for glycoproteins (Christianson et al., 2008) (Okuda-Shimizu & Hendershot 2007). BiP and PDI have been showed to be involved in ERAD by targeting a β-secretase isoform for degradation (Molinari et al., 2002)(. How and whether BiP and PDI can discriminate between folding intermediates and folding failures is unclear.

Although changes in local structure can be sufficient to retain a protein in the ESC, retention is not always this strict. Mutations in the ligand binding domain of the LDL receptor that cause hypercholesterolemia because of impaired LDL binding do not prevent the protein from leaving the ER and travelling to the cell surface (Hobbs et al., 1990). This is just one of many examples underscoring that quality control is based on structural and not functional criteria.

Very rarely glycoproteins are found to bind simultaneously to BiP and CNX or CRT. Therefore, it seems that a given glycoprotein enters first either the BiP or the CNX/CRT pathway. The initial choice is dictated by the localization of the N-glycans: the closer these are to the N-terminus of the nascent protein, the higher the tendency to use CNX as a chaperone

system (Molinari 2000). If the first attempts to fold fail, the protein can shift to the alternative pathway. Altogether, these data imply that sites of conjunction exist in which the substrate can jump from one pathway to another. In principle, however, it should be possible for large multi-domain proteins to engage with both.

1.3.1 Thiol-mediated retention

Studies of immunoglobulins (Ig) have revealed another mechanism of quality control, based on the recognition of exposed thiols on assembly intermediates (Thiol-Mediated Retention-TMR; Alberini et al., 1990). In plasma cells, IgM secretion is restricted to hexamers ($(\mu_2\lambda_2)_6$) or pentamers ($(\mu_2\lambda_2)_5$-J), in which individual subunits are linked by disulphide bonds involving a conserved cysteine residue (Cys575) in the IgM chain C-terminal tailpiece (μtp) (Fig. 1.12).

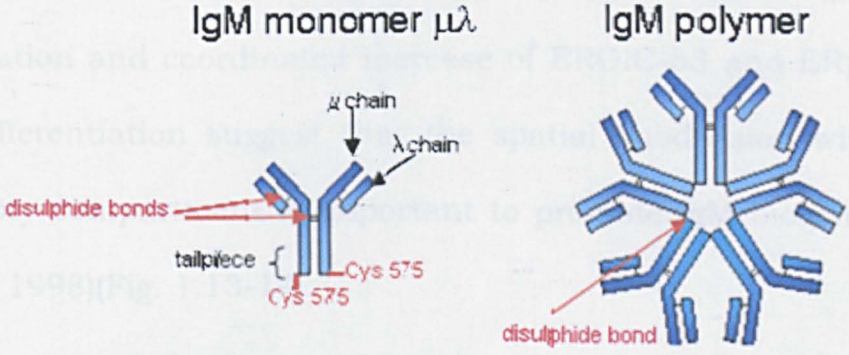


Figure 1.12. IgM structure.

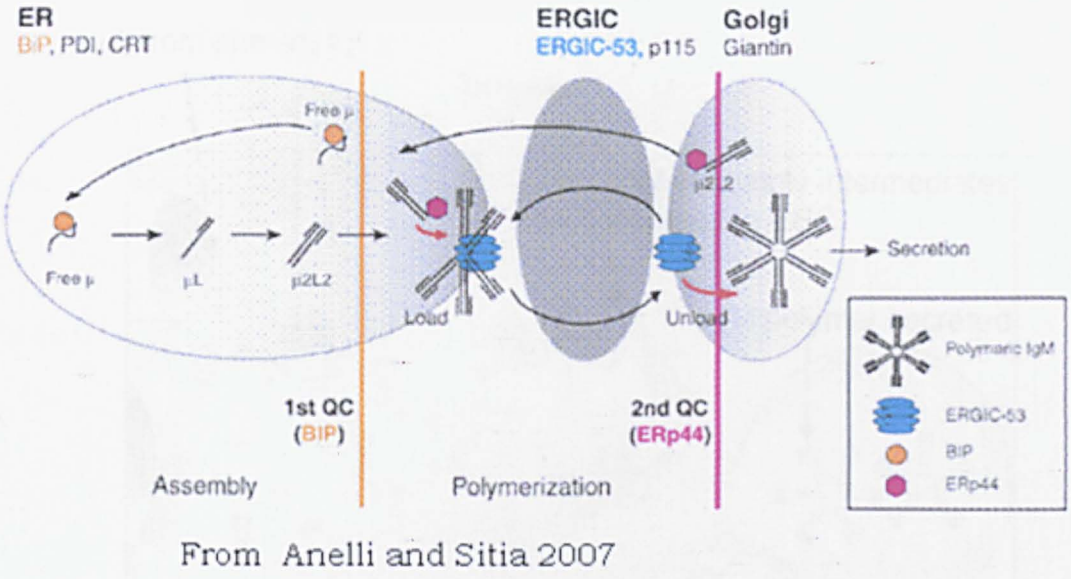
The building block of IgM, the "monomer" μλ. The μ and λ chains are disulfide bonded with one another. The μ tailpiece contains the Cys575 that is involved in IgM polymer assembly in pentamer or hexamer.

The first assembly step ($\mu_2\lambda_2$ formation) from individual μ and λ is fast and efficient in both B and plasma cells: its fidelity is mainly checked by BiP

(so called proximal/canonical QC, see Anelli and Sitia 2008) in the ER (Fig.1.13).

In addition, after a proximal QC, IgM seem to undergo also distal QC checkpoints in ERGIC and *cis*Golgi complex. Indeed, IgM polymerization is slow and occurs downstream to ER. Here a pool of ESC resident proteins (ERp44 and ERGIC-53-Anelli et al., 2007) support and check the oligomerization of unpolymerized IgM subunits that have already passed the BiP-dependent checkpoints (Anelli et al., 2003). As a hexameric membrane-embedded lectin, ERGIC-53 may provide a platform for IgM polymerization (Anelli et al., 2007). In the *cis*Golgi, ERp44 could capture unpolymerized IgM subunits and retrieve them via RDEL-dependent mechanisms.

The delayed interactions of nascent μ chain with ERp44 and ERGIC-53 with respect to BiP in B-lymphoma cells (Anelli et al., 2007) support the existence of sequential steps in IgM polymerization. The sub-cellular localization and coordinated increase of ERGIC-53 and ERp44 during B-cell differentiation suggest that the spatial subdivision within the early secretory compartment is important to promote IgM biogenesis (Reddy & Corley 1998)(Fig. 1.13-14).

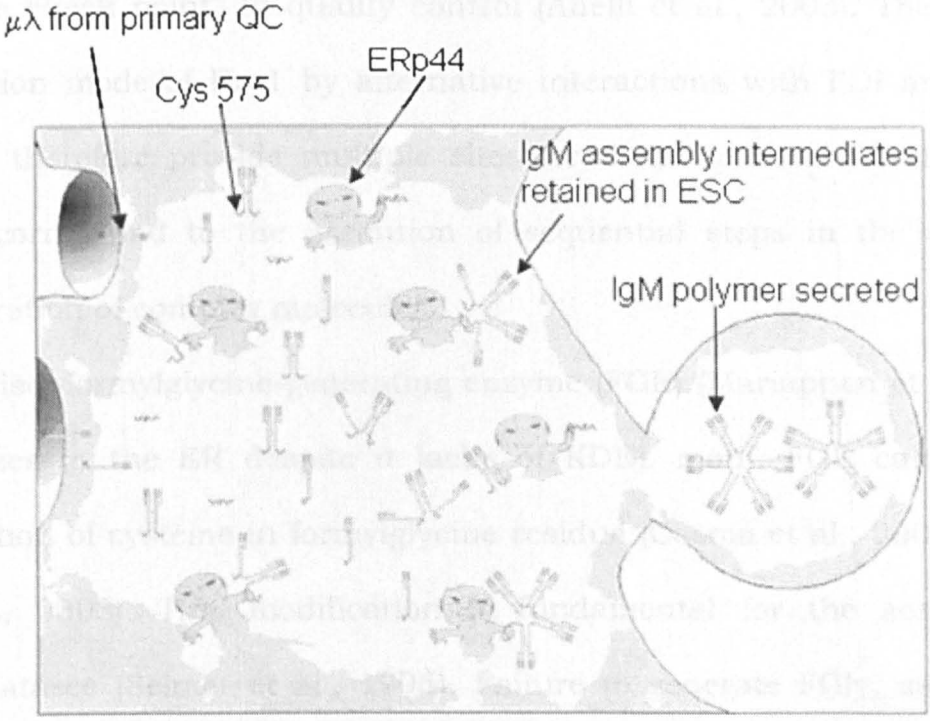


From Anelli and Sitia 2007

Figure 1.13. Schematic model of the IgM polymerization machinery

The IgM assembly occurs step-wise. The distribution of ERp44 in the early secretory pathway (primarily in the ERGIC) is depicted as a gradient of gray. BiP-dependent control ensures that μ chains do not proceed to the polymerization machinery unless assembled with λ chains (first QC step). ERGIC-53, a hexameric lectin captures $\mu_2\lambda_2$ subunits, likely aligning them in a planar conformation, suitable for polymerization. ERp44 ensures that unpolymerized subunits are not secreted, retrieving them into the assembly line (second QC step). Because of its ability to bind IgM subunits and ERGIC-53, ERp44 may recruit and locally concentrate $\mu_2\lambda_2$, to stimulate further the polymerization machinery efficacy.

Another example of sequential QC comes from adiponectin, a plasma protein secreted by adipocytes (Scherer et al., 1995) (Chandran et al., 2003). Plasma adiponectin is composed of trimers, hexamers and higher molecular weight oligomers (Bobbert et al., 2005)(Lara-castro et al., 2006). ERp44 was showed to retain folded adiponectin trimers, forming mixed disulfides with Cys39 in one subunit (Qiang et al., 2007) (Wang et al., 2007). As in the case of IgM, therefore, folded adiponectin intermediates are retained by thiol-dependent mechanisms.



Adapted from Reddy and Corley 1998

Figure 1.14. Schematic representation of the thiol-retention mechanism of IgM.

The thiol retention mechanism is illustrated by the presence of the "lumen trolls" (ERp44) whose function is to monitor the oxidation status of the Cys575 "hook" on all assembly intermediates. Assembly proceeds in a step-wise fashion using the subunits from this pool by means of the Cys575 of the μ heavy chains. Once the polymer has been covalently closed, Cys575 is masked and the quality control requirements of the ESC, therefore, allow the exit of the polymers.

Whilst the ER localization of soluble resident proteins largely depends on -KDEL receptor-mediated retrieval from the intermediate compartment or Golgi complex (Pelham and Munro, 1993)(Meunier et al., 2002), ER retention seems to be the fate of thiol-exposing proteins (Isidoro et al., 1996) (Anelli et al., 2003).

Considering that PDI and ERp44 can bind to different cysteines in Ero1α (Bertoli et al., 2004), the absence of ternary covalent complexes is surprising and may reflect temporal and/or spatial regulation within the

ER. ERp44 retains cargo proteins with exposed thiols, probably acting as a "late check point" in quality control (Anelli et al., 2003). The dynamic retention mode of Ero1 by alternative interactions with PDI and ERp44 could therefore provide multiple sites for disulfide bond formation that may correspond to the execution of sequential steps in the structural maturation of complex molecules.

Likewise, formylglycine-generating enzyme (FGE) (Mariappan et al., 2008) localizes in the ER despite it lacks of KDEL motif. FGE catalyzes the oxidation of cysteine in formylglycine residue (Cosma et al., 2003) (Dierks et al., 2003). This modification is fundamental for the activation of sulphatases (Selmer et al., 1996). Failure to generate FGly, as found in patients with mutations in the FGE-encoding SUMF1 gene, leads to multiple sulfatase deficiency (MSD), a rare but fatal inherited disorder that is characterized by synthesis of catalytically inactive sulfatases (Cosma et al., 2003) (Dierks et al., 2003). ERp44 interacts via Cys29 with Cys50 and/or Cys52 in the N-terminal domain of FGE. Von Figura and co-workers clearly showed that increasing ERp44 levels by over-expression improves and reducing ERp44 levels by silencing decreases the intracellular retention of FGE (Mariappann et al., 2008)

1.3.2 ERp44

In 2002, ERp44 was first identified as a protein involved in oxidative protein folding (Anelli et al., 2002).

ERp44 is a member of the PDI-family and is composed of an amino-terminal thioredoxin (TRX) domain a, endowed with a -CRFS- motif, followed by two redox inactive TRX-like domains, b and b', and contains a

carboxyterminal ER retrieval signal (-RDEL). Having a CRFS motif, ERp44 cannot act as an efficient oxidoreductase.

ERp44 homologues are present in *Mus musculus* (89% sequence identity), *Drosophila melanogaster* (51% identity), *Caenorhabditis elegans* (three hypothetical protein in which C30H7.2 displays a closest homology 45% identity). In *Saccharomyces cerevisiae*, the proteins with higher homology with ERp44 are Pdi1p and Eug1p, but identity is lower (~25%) and restricted to the TRX-like domain a.

ERp44 forms mixed disulphides via Cys29 (CRFS) in the Trx domain a (Anelli et al., 2003). When Cys29 is replaced with Ser, the formation of mixed disulphides with endogenous partners is prevented (Anelli et al., 2003). ERp44 contains six cysteines: mass spectrometry analyses revealed that the four C-terminal cysteines are covalently linked in sequential pairs (Anelli et al., 2002) whereas two are present as thiol (Cys29 and Cys63 in the domain a).

ERp44 contributes to retain Ero1α and Ero1β intracellularly, the two rate-limiting factors in PDI oxidation (Sitia et al., 2001), by forming mixed disulphides (Otsu et al., 2006). Moreover, ERp44 may play a role in the regulation of redox activity these proteins (Anelli et al., 2002). Similarly, ERp44 forms transient disulphides with IgM subunits (Anelli et al., 2002, 2003) (Anelli et al., 2007), adiponectin (Wang et al., 2007) or formylglycine-generating enzyme (FGE) (Mariappan et al., 2008), regulating their transport.

Recent evidence suggests that ERp44 is partly localized in ERGIC compartment and it is non-covalently bound to ERGIC-53 (Gilchrist et al., 2006) (Anelli et al., 2007). The localization of ERp44 also reflects its

capability of interacting with ERGIC-53. The ER-localized KKAA mutant of ERGIC-53 recruited endogenous ERp44 in this compartment. Nonetheless, some ERp44 is found in the ERGIC of the KKAA transfectants, which lack detectable ERGIC-53 in this compartment (Vollenweider 1998). By concentrating distally with respect to PDI, ERp44 and ERGIC-53 receive intermediates of assembly (IgM subunits) (Anelli et al., 2007) that have completed the first step in the production line, that is the BiP-dependent $\mu_2\lambda_2$ assembly (Hendershot and Kearney 1988)(Sitia et al, 1990); supporting the oligomerization within ESC.

Consistent with its role in folding and transport in the secretory pathway (Anelli et al, 2002) (Anelli et al., 2003) (Anelli et al., 2007), ERp44 is induced during ER stress and abundantly expressed in secretory tissue cells (Human Protein Atlas; www.hpr.se).

In addition to its crucial role in thiol-mediated retention and quality control, ERp44 interacts with inositol 1,4,5-trisphosphatereceptor type I, inhibiting its Calcium channel activity (Higo et al., 2005) (see 1.4).

1.3.3 ERGIC-53

ERGIC-53 belong to a class of type I transmembrane sorting receptors (Appenzeller-Herzog & Hauri 2006) and consist of an L-type lectin domain, a stalk domain, a transmembrane domain, and a short cytoplasmic domain. ERGIC-53 forms disulfide-linked homodimers and homohexamers through 2 cysteine residues in the stalk domain. (Lahtinen et al., 1999) ERGIC-53 mainly localized in ERGIC and is continuously shuttled from ER to ERGIC. Indeed, ERGIC-53 has a C-terminal KKFF di-lysine ER-targeting signal that binds to COP I proteins and is required for

recycling. Two C-terminal phenylalanine residues in the short cytosolic domain are also necessary for the proper trafficking of ERGIC-53 from the ER (Nufer et al., 2002). ERGIC-53 traffics between the ER, where it captures soluble glycoproteins, and the Golgi complex, where it releases them thanks to lower Calcium concentration and pH (Appenzeller-Herzog et al., 2004). These two factors in fact regulate both the protonation of a conserved histidine in the ERGIC–53 lectin domain and its interaction with other regulatory proteins, such as MCFD2 (Hauri et al., 2000) (Appenzeller-Herzog et al., 2004) (Kawasaki et al. 2008). Its cargo is a specific subset of glycoproteins, including blood coagulation factors V and VIII (Zhang et al. 2003), two lysosomal enzymes glycoproteins cathepsin Z (Appenzeller et al., 1999) and cathepsin C (Vollenweider 1998) (Nyfeler et al., 2006), procathepsin Z (Appenzeller-Herzog & Hauri 2006), and alpha 1-antitrypsin (Nyfeler et al. 2008).. Besides, ERGIC-53 is also involved in the traffic of IgM (Mattioli et al. 2006) (Anelli et al., 2007), and Sulfatase modifying factor 1(Fraldi et al. 2008).

Sitia and co-workers showed the concerted action of ERp44 and ERGIC-53 in IgM polymerization (Anelli et al., 2007): hexameric ERGIC-53 may provide a planar platform that concentrates $\mu_2\lambda_2$ subunits and favours their ordered assembly (Fig. 1.13), avoiding formation of larger polymers (Lalla & Sitia, 1998). In support of this model, the inactive N156A mutant inhibits polymerization, likely because hetero-hexamers (with endogenous wt molecules) with reduced valency are formed (Anelli et al., 2007).

Consistently with its role in regulating the transport of coagulation factors V and VIII, the gene that encodes ERGIC-53 has been identified as a gene

responsible for the autosomal recessive bleeding disorder combined deficiency of coagulation factors V and VIII (F5F8D)(see 1.5).

1.4 ESC and signalling

Often, intracellular signalling is triggered by an external ligand targeting a Plasma Membrane (PM) receptor. The signal is then transduced into and propagated through the cell by a well-characterized cascade that involves specific protein categories, such as kinases, phosphatases, lipases and GTPases. Over the last decade, several studies reported the presence of these signalling proteins on different intracellular organelles, including endosomes, ER, Golgi complex, mitochondria, and the nuclear membrane (Sallese et al., 2009).

The ESC is indeed a crossroads of at least three signalling pathways: signalling from KDEL receptor, unfolded protein response and Calcium signalling.

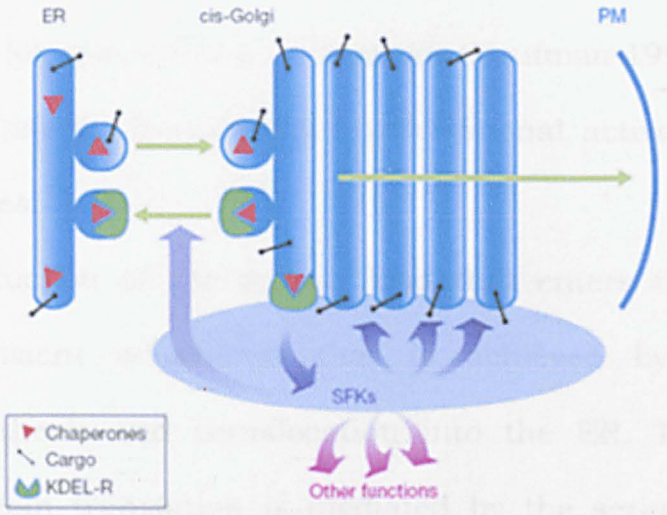
1.4.1 Signalling from KDEL receptor

Sallese and co-workers studies (Pulvirenti et al., 2008) uncovered a signalling reaction that is initiated by a Golgi-localized signal–receptor system that senses, and is activated by transport carriers arriving at the Golgi complex from the ER. This triggering signal is provided by the ESC chaperones. As these chaperones reach and enter the GC they bind to and 'activate' the KDEL receptor (KDEL-R) through their KDEL motif (Munro Pelham 1987). The KDEL-R is a seven-transmembrane domain-containing protein that once activated, dimerizes, changes conformation within the

membrane, and recruits the COPI trafficking machinery. These then move retrogradely, thus recycling the chaperones to the ER, from where the KDEL-R can again shuttle between the ER and the Golgi complex. KDEL-R binding to ESC chaperones could initiate activation of the Src family kinases (SFKs), and hence a SFK-dependent tyrosine phosphorylation cascade on the Golgi complex (Pulvitrenti et al., 2008). In turn, this activates the intra-Golgi transport machinery, and thus balances the membrane input to the Golgi that initiates the whole process with an increased membrane output from the Golgi complex. Although, the way in which the exit of the chaperones from the ER is linked with, or regulated by, the efflux of cargo from the ER remains to be determined, once in the Golgi complex, the chaperones can activate a KDEL-R-dependent phosphorylation cascade that controls intra-Golgi trafficking.

At its most basic level, this chaperone/KDEL-R/SFK mechanism provides a simple feed-forward circuit that can result in positive gating of the Golgi complex, and that helps to preserve the dynamic equilibrium of the Golgi complex during trafficking

This chaperone/KDEL-R/SFK system might be a 'universal' traffic-regulatory device. This is suggested by the ubiquitous nature of chaperones, which would allow them to serve as general trafficking signals in all cell types, and is further supported by the dependence of membrane trafficking on the SFKs that has been seen in all cell lines tested to date (Pulvirenti et al., 2008) (Fig. 1.15).



From Pulvirenti et al., 2008

Figure 1.15. Model of traffic self-regulation by the chaperone-KDEL-R-SFK signal response system.

Chaperones containing KDEL sequences (red triangles) leave the ER with cargo (black sticks). On arrival at the *cis*-Golgi, they bind the KDEL-R (green Pacmen), triggering activation of SFKs on the Golgi complex. This activation may occur by direct interactions between the KDEL-R and Golgi SFKs or by a more complex mechanism. In turn, these p-SFKs induce a phosphorylation cascade (violet arrows) that is required for the trafficking machinery to function for the subsequent intra-Golgi transport.

1.4.2 Unfolded protein response

ER stress can be induced by a variety of physiological conditions, including perturbations in calcium homeostasis, glucose/energy deprivation, redox changes, ischemia, viral infections and protein mutations, is alleviated by the Unfolded Protein Response (UPR) (Kaufman 2002) (Szegezdi et al., 2006)(Fig. 1.16).

Perturbations that alter ESC homeostasis might disrupt folding and lead to the accumulation of unfolded proteins and protein aggregates, which are detrimental to cell survival.

On a cellular level, the adaptive phase of the UPR triggers three kinds of protective cellular responses:

(i) Up-regulation of ESC chaperones such as BiP/GRP78 to assist in the folding-assembly of proteins (Kaufman 1999) is a longer-term adaptation that entails transcriptional activation of UPR target genes;

(ii) Reduction of the protein load that enters the ER, which is a transient adaptation that is achieved by lowering protein synthesis and translocation into the ER. This attenuation of protein translation is mediated by the serine–threonine kinase PERK (protein kinase-like endoplasmic reticulum kinase) which phosphorylates the initiation factor-eIF2a thereby reducing translation (Prostko et al.,1993);

(iii) Degradation of misfolded proteins by the proteasome by a process called ER associated degradation (ERAD) (Hampton, 2002)

ESC triggered signalling pathways through three proximal sensors of the unfolded protein response, inositol requiring element-1 (IRE-1), PKR-like ER kinase (PERK) and activating transcription factor 6 (ATF6) in the ER. Excessive and/prolonged stress leads to a maladaptive response and apoptosis (Szegezdi et al., 2006).

ATF6, PERK and IRE1 are integral membrane proteins constituted by a luminal portion that sense the protein-folding environment in the ER, and by a cytoplasmic effector portion that interact with the transcriptional or translational apparatus in the cytosol and in the nucleus.

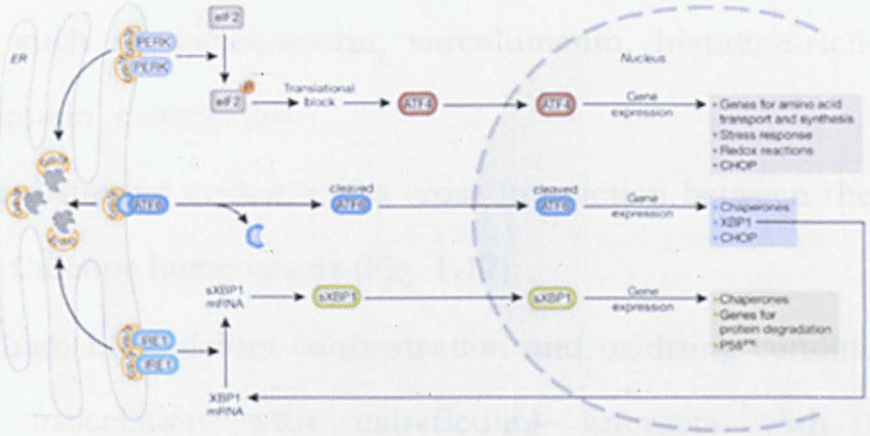
All three of these sensors are maintained in an inactive state at the ER membrane by binding to BiP. Upon accumulation of unfolded proteins,

bound BiP dissociates from ATF6, IRE-1 and PERK to chaperone the misfolded proteins thereby permitting the activation of one or more of these transducers (Szegezdi et al., 2006). BiP therefore serves as the master regulator of the UPR.

PERK is a type I transmembrane serine threonine kinase. When activated, PERK homodimerizes and self-phosphorylates, thereby activating itself and initiating its eIF2a kinase activity. Phosphorylation of the translation initiation factor, eIF2a, results in the formation of a stalled 43S ternary complex that causes a general decrease in the translation of most proteins. However some selected proteins with internal ribosomal entry sites (IRES), such as ATF4, GRP94 and BiP, are translated more efficiently (Harding et al., 2000) and hence their protein levels actually increase.

Once activated, the cytoplasmic domain of IRE-1a gains endoribonuclease activity and cleaves 26 nucleotides from the mRNA encoding X-box binding protein (XBP) 1, generating a spliced variant (XBP-1s) that functions as a potent transcriptional transactivator of genes involved in ER expansion, protein maturation, folding and export from the ER, as well as export and degradation of misfolded proteins (Yoshida et al., 2001) (Calfon et al., 2002)(Yoshida et al., 2003). IRE-1 may also degrade ER-targeted mRNAs, thus decreasing the production of new proteins in the organelle (Hollien & Weissman 2006) (Pirot et al., 2007). ATF6 signalling and transcription of ER chaperones ATF6 is a 90-kDa-bZIP protein that is activated by post-translational modifications. ATF6 activation as part of the UPR leads to its translocation to the Golgi and cleavage by site-1 protease (S1P) and S2P. The 50-kDa cleaved ATF6a translocates to the cell nucleus, where it binds to the ER stress response element CCAAT (N)

9CCACG (Yoshida et al., 1998) in genes encoding ER chaperone proteins such as BiP and GRP94. This results in increases in the level of these proteins and hence increased protein folding activity in the ER (Okada et al., 2002).



From Szegezdi et al., 2006

Figure 1.16. The unfolded protein response

On aggregation of unfolded proteins, GRP78 dissociates from the three ER stress receptors, PERK, ATF6 and IRE1, allowing their. This concerted action aims to restore ESC function by blocking further build-up of client proteins, enhancing the folding capacity and initiating degradation of protein aggregates.

1.4.2 Calcium signalling from ER

The ER is the most important intracellular Calcium store that can accumulate it to concentrations of 10–100 mM, while its concentration in the cytoplasm of the resting cell remains within the range of 100–300 nM (Görlach et al., 2006). Upon stimulation of plasma membrane (PM) receptors or upon electrical excitation of the PM, the ER releases Calcium, thus participating in the generation of rapid Calcium signals. Since the ER storage capacity is limited, Calcium release must be followed by Calcium replenishment. Calcium movements across the ER membrane are

facilitated by three classes of proteins: Calcium release channels inositol-1,4,5-triphosphate (IP_3) receptors (IP_3Rs) (Bezprozvanny 2005)(Mikoshiba 2007) and ryanodine receptors (RyRs) (Hamilton 2005), Calcium re-uptake pumps–sarco-endoplasmic reticulum Calcium-ATPases (SERCAs) (Periasamy & Kalyanasundaram 2007) and luminal Calcium-binding proteins, such as calsequestrin, sarcalumenin, histidine-rich Calcium-binding protein, calreticulin.

There is substantial evidence of a cross interaction between the ER redox state and Calcium homeostasis (Fig. 1.17):

- At high ER Calcium concentration and oxidizing condition, ERp57 (in association with calreticulin) interacts with SERCA2b, inactivating the pump by oxidation, while RyRs and IP_3R1 are open. (Li & Camacho 2004).
- At low ER Calcium concentration and reducing conditions, ERp44 directly inhibits the channel IP_3R1 (Higo et al., 2005) and calsequestrin (CSQ) inhibits the channel RyR (Szegedi et al., 1999).

It has been reported that ERp44 can directly interact with the third luminal loop (L3V) of IP_3R type 1 (IP_3R1) to inhibit its channel activity in a pH-, redox state and ER Calcium concentration dependent way (Higo et al., 2005). Indeed, a short region of ERp44 rich in glutamic acid residues (from 207 to 256), which shows no sequence similarities with other known proteins but is well-conserved among species, was reported to be necessary and sufficient for the interaction with the IP_3R1.

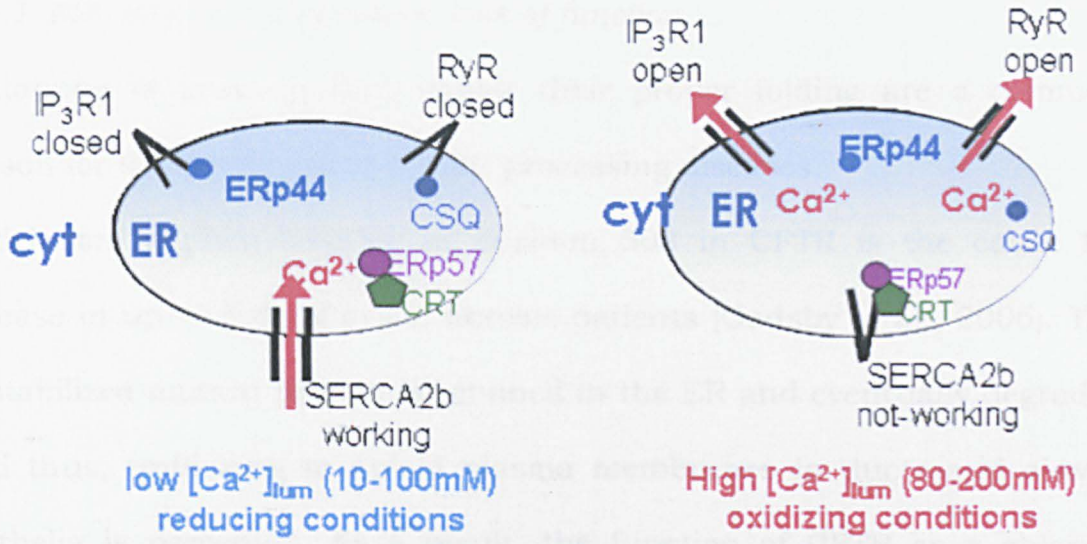


Figure 1.17. Overview of cross interaction between the ER redox state and Calcium homeostasis.

CRT: Calreticulin; RyR: Ryonidin receptor; Cyt: cytosol; CSQ: Calsequestrin. See text for details

1.5 ESC processing diseases

Little is known about the overall efficiency of protein folding and assembly in the ESC. A number of examples describe a wide range from very high efficiency observed with viral proteins (e.g. vesicular stomatitis virus glycoprotein is about 90%-(Mathieu et al., 1996) to very low efficiency for ion channels (cystic fibrosis transmembrane conductance regulator, CFTR- about 20%-(Kopito et al., 1999)

Likewise, overall protein production in the ESC may not be of high efficiency and dedicated machinery is built into the folding and assembly machine of the ESC to degrade/remove non-native proteins and thus preventing intracellular protein accumulation and aggregation.

1.5.1 ESC processing diseases: loss of function

Mutations in proteins that impair their proper folding are a common reason for the development of ESC processing diseases.

A deletion of phenylalanine at position 508 in CFTR is the cause for disease in up to 80% of cystic fibrosis patients (Gadsby et al., 2006). The destabilized mutant protein is retained in the ER and eventually degraded and thus, trafficking to apical plasma membranes in ducts and airway epithelia is prevented. As a result, the function of CFTR as a chloride channel is compromised. Loss of CFTR function alters the hydration of the mucosal layer that lines the airway epithelia leading to persistent bacterial growth and development of lung fibrosis (Rowe et al., 2005).

In α1 antitrypsin (α1AT) deficiency, mutant α1AT Z cannot be secreted by hepatocytes into the blood and body fluids, where it normally inhibits neutrophil proteases. As a consequence, proteases degrade the connective tissue matrix of the lung causing emphysema (Gooptu & Lomas 2009). This is a clear case of overzealous QC, in that α1AT Z retains enzymatic activity but cannot deploy it owing to intracellular retention.

Likewise, loss of function derived from folding problems in a variety of proteins leads to retention of mutated proteins and their degradation (Aridor 2007).

1.5.2. ESC processing diseases: Gain of function

In addition to loss-of-function derived from the inability of such protein to be sorted in its final destination, a gain-of-function occurs when the ESC is unable to dispose the mutant proteins. Many human diseases are

characterized by the presence of protein aggregates in the lumen of the ESC sub-regions. As a result, it may lead to toxicity and cell death.

Hereditary hemochromatosis (HH) is a recessive disorder characterized by pathological abnormal iron accumulation in tissues. The pathology is due to a cysteine to tyrosine substitution in HFE folding and association with β2 microglobulin (Gray et al., 2009). Impaired assembly prevents surface expression of HFE, the majority of the mutant protein remaining in the Esc, as high molecular weight aggregates (Gray et al., 2009).

Abnormal plasma cells (Mott cells) displaying dilated, Ig-containing ER cisternae (Russell Bodies, RB) are present in some non-secretory myelomas and in autoimmune or inflammatory diseases (Jiang et al. 1997). RB form when mutant Ig is produced that can neither be secreted nor degraded (Kopito & Sitia 2000) (Mattioli et al., 2006).

In the absence of λ chains, μΔCH1 (μ lacking the BiP-binding CH1 domain) condenses in ERGIC-derived tubular structures (smooth RB, sRB). In its presence, larger, ribosome-coated vesicles (rough RB, rRB (Mattioli et al., 2006) are formed. rRB and sRB likely reflect abnormalities in the proximal and distal QC steps assisting IgM polymerization.

The observation of Mott myelomas with RB and the capability of selecting B cells with aggregation imply that cells can coexist and proliferate with abundant aggregates in their ESC. Why do other aggregates cause disease? A possible explanation is that cell division generates space for accommodating aggregate-containing vesicles. In a tissue where instead proliferation is limited by space constraints, the accumulation of dilated ESC cisternae is instead bound to cause problems to the cell and its neighbours.

1.5.3 ESC processing diseases: machinery diseases

Mutation in ESC-protein folding crew can also lead to ESC-derived diseases. Highly specific defects develop as observed in cranio-lenticulo-sutural dysplasia (CLSD) (Boyadjiev et al., 2006). Orci and co-workers showed that CLSD occurs as a result of defective COPII-mediated endoplasmic reticulum export owing to loss of function of SEC23A. As a result, collagen and (probably) other secretory proteins accumulate and enlarge the ER, ultimately leading to the clinical manifestations of CLSD. The relatively mild phenotype of affected individuals suggests that the 1144T-C SEC23A mutation is a hypomorph and that the mutant protein retains some residual functional activity (Boyadjiev et al., 2006.)

Mutations that disrupt the SIL1 function (also called BAP) protein cause Marinesco-Sjögren syndrome (MSS) (Senderek et al. 2005), a form of a cerebellar ataxia complicated by cataracts, developmental delay and myopathy. SIL1 acts as a nucleotide exchange factor for chaperone BiP, which is a key regulator of the main functions of the ER.

Loss-of function mutations in ERGIC-53 protein cause a recessive bleeding disorder characterised by decreased levels of two glycoproteins (Combined Deficiency of Coagulation Factors 5 and 8, F5F8D) (Zhang et al. 2003) (Spreafico & Peyvandi 2008). In other F5F8D patients, mutations are localized in MCFD2, a protein that associates with ERGIC-53 (Zhang et al., 2003). Despite that ERGIC-53 is involved in many processes including IgM polymerization, SUMF1 folding/transport and α1AT traffic, only Factors 5 and 8 traffic is crucially dependent by ERGIC-53

suggesting compensatory mechanisms and/or redundancy in the ESC quality control crew (VIPL and VIP36).

2.Aims of the work

ERp44 plays a key role in Thiol-Mediated Retention (TMR) of variety of client proteins (Ero1α-Anelli et al., 2002; Anelli et al., 2003. SUMF1-Mariappan et al., 2008; Fraldi et al., 2008. Adiponectin-Wang et al., 2007. IgM-Anelli et al., 2007; Cortini & Sitia 2010) and thus, regulating their transport and localization. Besides being fundamental to the quality control system, ERp44 interacts with IP_3R1 (Inositol trisphosphate receptor 1), inhibiting its activity and thus, regulating the efflux of calcium ions from ER (Higo et al., 2005). Altogether these features make the ERp44 a candidate to integrate transport protein signalling and redox homeostasis.

Several open questions remain to be investigated:

I. What is the three-dimensional structure of ERp44?
II. How does monomeric ERp44 find the oxidative power to form disulfide bonds with its substrates?
III. Which is the molecular basis of TMR of ERp44?
IV. How is ERp44 regulated in living cells?

The overall aim of this work is to obtain more insights into the mechanisms of ERp44-TMR, by means of structural and functional approaches.

First of all, a recombinant form of human ERp44 was produced and purified to homogeneity. This material was used for crystallization studies aimed at the determination of the three-dimensional structure of ERp44

using X-ray crystallography. The availability of high-resolution structure of a protein is mandatory for understanding its structure-function relationship. It is particularly important to assess at the atomic level the role of the individual domains of ERp44 in modulating its diverse functions. Moreover, three-dimensional structure provided a solid framework for the design of protein mutants of ERp44.

Exploiting a combination of different approaches, structure-based functional analysis aimed to elucidate the molecular mechanisms of ERp44 regulation. While X-ray crystallography could produce valuable high-resolution information on stable conformations of macromolecules, it cannot readily provide information on dynamics. Thus, we used a combination of *in silico* mutagenesis and molecular dynamics (MD) to gain insights into the role of specific amino acid residues of the dynamic regulation of ERp44.

In addition, ERp44 mutants devised from both the structure and *in silico* experiments were extensively characterized. ERp44-mediated Ero1α retention/interaction was exploited as main assay to assess the role of ESC pH on ERp44 physiology and to determine the role of individual domain/region/residues in TMR activity.

3. Results

3.1 ERp44 production, purification and crystallization

3.1.1 Protein production in *E.coli*

A preliminary bioinformatic analysis of human ERp44 protein sequence, differently from what previously reported (Anelli et al., 2002) (Anelli et al., 2003)(Higo et al., 2005), revealed that it contains a N-terminal TRX-like domain, similar to that of PDI a, with a –CRFS- motif followed by two redox-inactive TRX-like domains, homologous to the b and b' domains of PDI (Fig. 3.1).

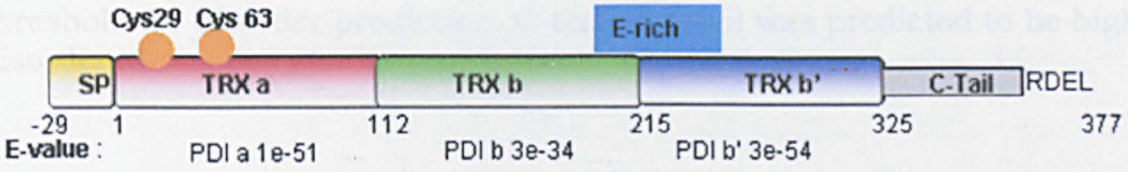


Figure 3.1 ERp44 domain organization

Bioinformatic analysis of human ERp44 protein sequence by Conserved Domain Database (CDD) (Marchler-Bauer et al., 2005) reveals that it contains a three TRX-like domain, similar to PDI (see E-value). Orange balls are cys residues in domain a. Cys 29 mediates covalent interaction with client proteins (Anelli et al., 2002). The three thioredoxin-like domains are followed by a disordered C-Tail of 52 residues (from 325 to 373) and by –RDEL motif for ESC retrieval. The Signal Peptide (yellow box) is cleaved upon entry in the ER, yielding the mature form of ERp44 (1-377). E-rich region, depicted in blue, mediates the interaction with IP_3R1.

The three thioredoxin-like domains are followed by a disordered C-Tail of 52 residues (from 325 to 377 residues) (Fig. 3.2) with no obvious sequence homology with other proteins (Fig. 3.1) The cDNA fragments encoding the full-length ERp44 (without –RDEL C-terminal motif and N-terminal signal peptide) was cloned into a pET-28b vector, allowing the expression of a

hexahistidine tag fused to the N-terminus of the recombinant protein for expression in *E. coli* (Fig. 3.3).

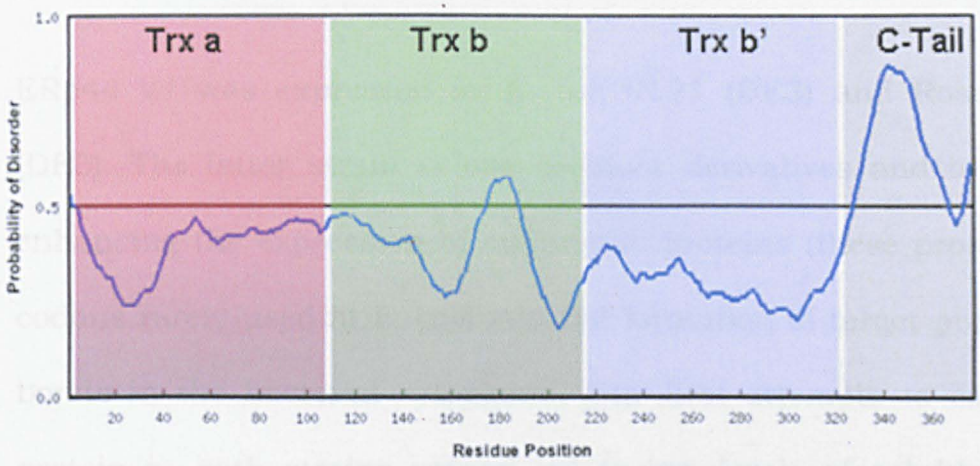


Figure 3.2. Probability of disorder for human ERp44

RONN (Regional Order Neural Network) plot (Yang et al., 2005) of disorder probability per residue for ERp44. The horizontal dashed lines mark the threshold for disorder prediction. C-terminal tail was predicted to be highly disordered.

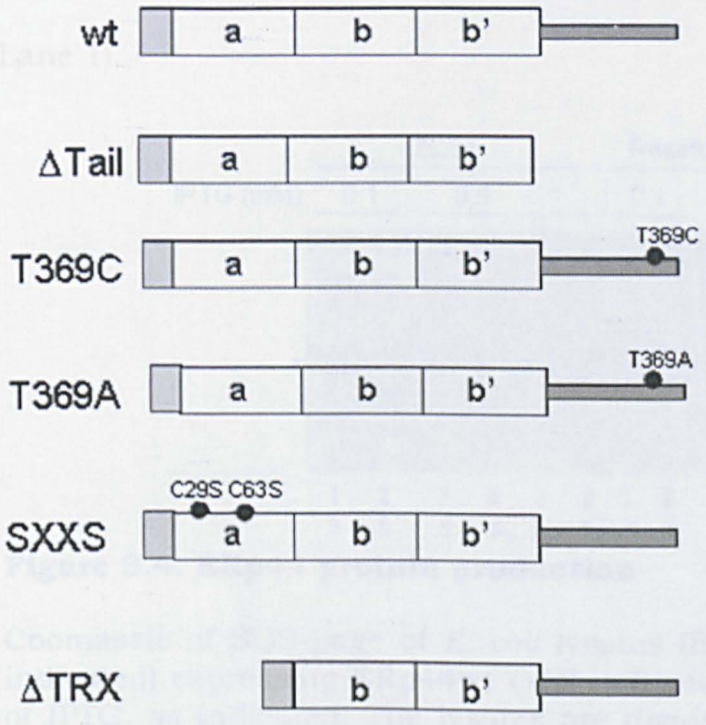


Figure 3.3. ERp44 protein produced in *E. coli*

Schematic representation of ERp44 cloned in pET28b vector for *E. coli* expression. Grey square at N-terminus represents His-tag. Black circles indicated point mutations.

Among the six cysteine residues which ERp44 consists, four cysteine residues form two disulfide bonds and the other two sulfhydryls are predicted as free thiols (Anelli et al., 2002).

ERp44 wt was expressed in *E. coli* BL21 (DE3) and Rosetta-gami 2™ (DE3). The latter strain is one of BL21 derivatives and is designed for enhancing the expression of eukaryotic proteins (these proteins contains codons rarely used in *E. coli*) and the formation of target protein disulfide bonds in the bacterial cytoplasm. Our first attempts to express ERp44 protein in both strains indeed led to low levels of soluble protein (not showed). Lowering of the IPTG concentration to 0.1 mM and changing the induction temperature to 20°C resulted in extremely high yields on soluble protein (Fig. 3.4, lanes 1 and 2). BL21 (DE3) produced more recombinant protein than Rosetta-gami 2™ (DE3), and thus, BL21 (DE3) was chosen for the following large scale expression of recombinant ERp44 (Fig. 3.4-Lane 1).

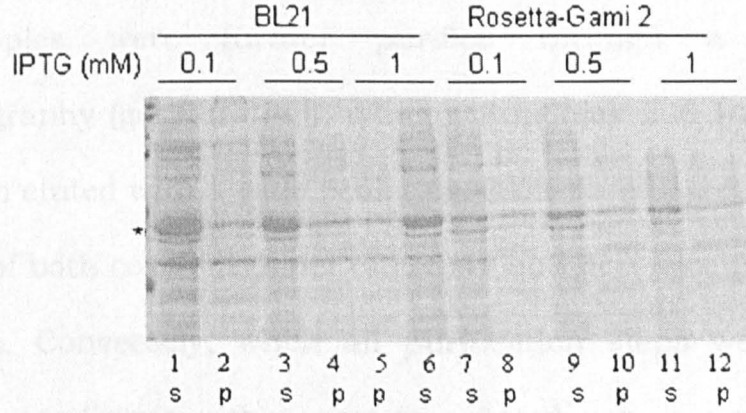


Figure 3.4. ERp44 protein production

Coomassie of SDS-page of *E. coli* lysates (BL21 and Rosetta-Gami 2, as indicated) expressing ERp44wt O/N induced with different concentration of IPTG, as indicated. The lysates are divided in soluble fraction (S) and insoluble fraction (P). The insoluble fraction is resuspended in SDS-page loading buffer. Asterisk indicated His-tag ERp44.

3.1.2 Purification of recombinant His-ERp44

The pET28-ERp44 plasmid was used to transform *E. coli* BL21(DE3) to obtain an amount of purified recombinant protein sufficient for extensive crystallographic characterization. The purification of ERp44 was achieved through sequential chromatographic steps. Initially, the polyhistidine tag was utilized for affinity chromatography using Ni-NTA resin. After elution, all the collected fractions were analyzed by SDS-page before pooling them together. A high degree of purity was achieved in this first purification step, as confirmed by the presence of a single well defined band at about 45 kDa when purification is performed in reducing condition (1 mM 2ME) (Fig. 3.5a- panel "Ni-NTA red"). Indeed, several bands co-purify with ERp44 when Ni-NTA purification was performed under non-reducing conditions, indicating the formation of covalent complexes between ERp44 and different endogenous proteins in *E. coli* (Fig. 3.5a- panel "Ni-NTA Nred").

The samples were further purified through a size exclusion chromatography (gel filtration), using a Superdex 200 HR 10/30 column, the protein eluted with a wide peak (Fig. 3.5b- in blue), consistent with the presence of both covalent dimer of ERp44 and non-specific mixed disulfide complexes. Conversely, when all purification steps were performed in reducing conditions, the protein eluted at a retention volume corresponding to a molecular weight of 50 kDa consistently with a monomeric protein (Fig. 3.5b in red). The recovered recombinant protein was incubated with thrombin and the quantitative removal of the polyhistidine tag was assessed by Coomassie staining and protein blot

with the anti-His probe (Fig. 3.5c). A further size exclusion chromatography was necessary to separate thrombin, the fusion tags, and any possible product of degradation from the recombinant protein of interest. The final yield of ERp44 was 15 mg of pure protein per litre of bacterial culture. The sample was 95% pure as estimated by SDS–page (Fig. 3.5a panel "Post-FPLC Red").

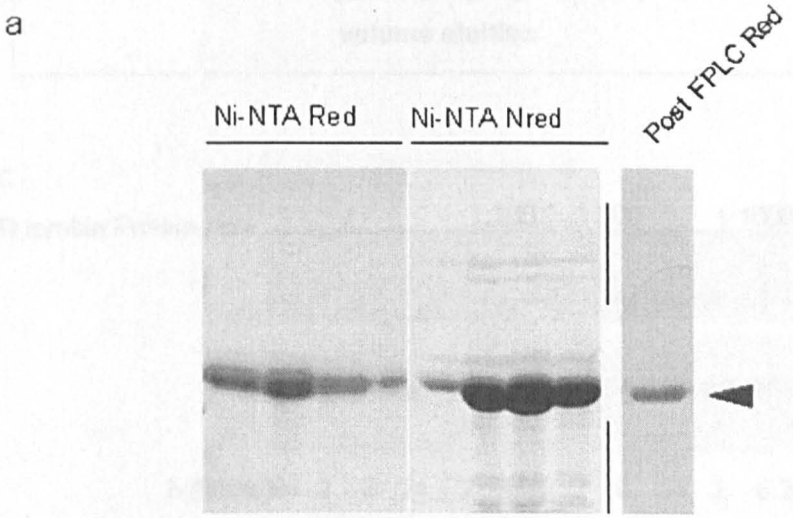
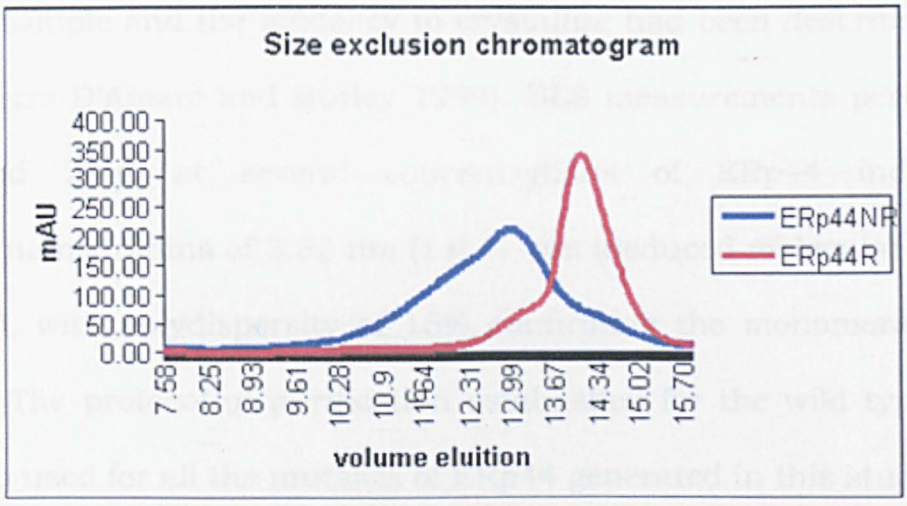


Figure 3.5a. ERp44 purification

A. NTA ERp44 protein purification. Coomassie SDS-page of elution fractions of a Ni-NTA purification of ERp44, performed in reducing (R) or non-reducing (Nred) conditions, as indicated. The lane labelled as "Post FPLC Red" shows the purity of the protein after size exclusion chromatography.Blue arrowhead indicates ERp44

b



c

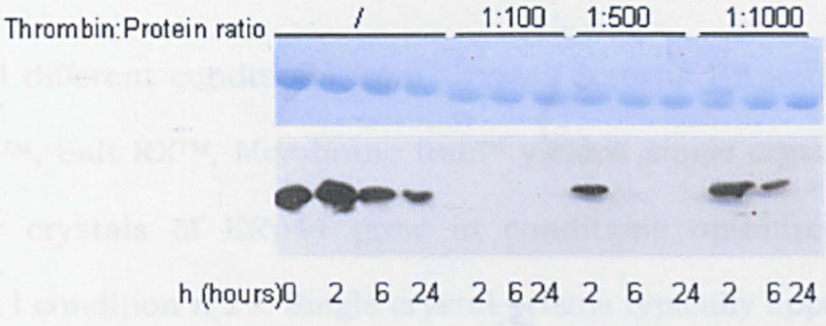


Figure 3.5b and c. ERp44 purification

B. Size exclusion chromatography ERp44 purification. Size exclusion chromatograms of pooled Ni-NTA purification fraction performed in reducing condition (R; red line) or non-reducing condition (NR; blue line).

C. Removal of His-tag from ERp44 fusion protein with thrombin. Combination of time and Thrombin/protein ratio monitored by Coomassie (upper panel) and protein blot (α-His tag-bottom panel). Upon His-tag removal, ERp44 undergoes a small MW shift in Coomassie and is not detected anymore by α-his antibodies in protein blots. The used protocol for large scale cleavage is 24 h with 1:1000 thrombin:protein ratio.

Dynamic Light Scattering (DLS) is widely used to assess the monodispersity of the sample. Indeed, the level of polydispersity determinated by DLS is related to the likelihood of crystallization: a

positive correlation of low (generally ≤ 20 %) polydispersity of a given protein sample and the tendency to crystallize had been described (D'arcy 1994)(Ferre-D'Amare and Burley 1994). DLS measurements performed at 4°C and 20°C at several concentrations of ERp44 indicated a hydrodynamic radius of 3.52 nm (± 0.1) nm (deduced molecular weight of 60 kDa), with polydispersity of 15% confirming the monomeric state of ERp44. The protocol of purification established for the wild type protein was also used for all the mutants of ERp44 generated in this study.

3.1.3 Crystallization, diffraction improving and data collection

Several different conditions from Crystal Screens I™ and II™ and Index Screen™, Salt RX™, Membrane frac™ yielded single crystals. Diffraction-quality crystals of ERp44 grew in conditions optimized from Crystal Screen I condition n°29. Single crystal prisms typically appeared after 48 h in 0.85 M Na/K Tartrate 0.1 M Tris-Cl pH 8.5 and grew to maximum dimensions of 0.3 x 0.3 x 0.5 mm in 4 days (Fig. 3.6).

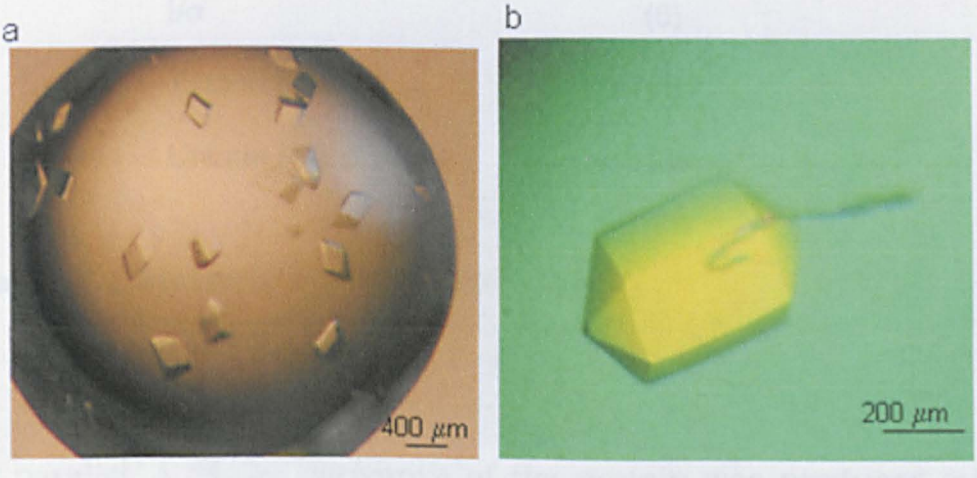


Figure 3.6 ERp44 crystals

Pictures of full length ERp44 crystals (A-his-tagged ERp44. B-native ERp44).

The native crystals diffracted to 5.0 Å in house. One native data set was collected to 4.0 Å at the beam line ID29 of the European Synchrotron Radiation Facility in Grenoble (France) (Fig. 3.7). These crystals belong to the trigonal space group $P3_121$ or $P3_221$, with unit-cell parameters a = 88.36, b = 88.36, c = 135.04 Å and α= 90°, β= 90°, γ= 120°. The unit cell volume is consistent with one molecule of ERp44 in the asymmetric unit (Table 3.1).

Data collection	
Resolution range (Å)	50-4
Wavelength (Å)	0.97372
Space group	P321
Cell dimension	
a,b,c(Å)	88.36, 88.36,135.06
α,β,γ (°)	90,90,120
Completeness (%)	93.8(69.4)
R_{sym} (%)	9.7(40)
I/σ	(6)
Total reflections	42761
Unique reflections	5139

Table 3.1 Statistic for ERp44 His-tag

In parallel, a SeMet derivative of the protein was produced following the method of Liu et al., (Liu et al., 2005) Quantitative SeMet incorporation was confirmed by mass spectrometry (data not-showed). Although, the SeMet preparation yielded crystals under very similar conditions as the

wild type protein, but no SAD (single-wavelength anomalous dispersion) data set could be collected since crystals of the SeMet-labeled ERp44 did not diffract to high resolution.

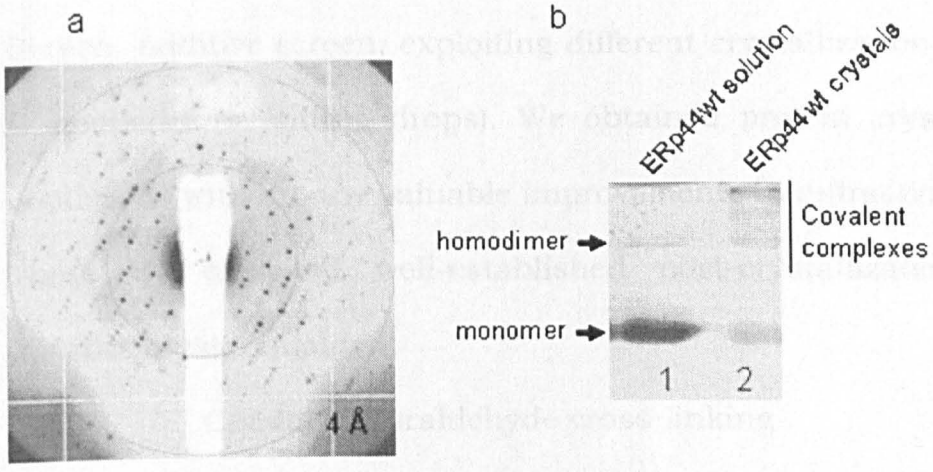


Figure 3.7. ERp44 diffraction analysis

A. Diffraction pattern of the ERp44 crystals which diffracted to barely beyond 4 Å resolution (ESRF ID-29).

B. ERp44 crystal SDS-page non-reducing SDS-page coomassie of protein crystals and protein in solution as indicated. Arrows indicate monomeric and homodimeric protein while disulfide complexes are indicated with bracket.

SDS-page in non-reducing conditions of dissolved ERp44 crystals revealed that the protein, during crystallization process, partially forms disulfide complexes (Fig. 3.7b- lane 2). Indeed, this could create micro heterogeneity in the crystal lattice and therefore affects the diffraction quality. DLS analysis revealed that after 3-4 days of storing at 4° (not-showed), monomer ERp44 partially forms homodimers confirming that ERp44 is reactive also in solution (Fig. 3.7b- Lane 1). To overcome this situation, we added 5 mM DTT before crystallization but, unfortunately, this was not enough to keep ERp44 reduced along equilibration period.

Three strategies were exploited to improve crystal quality and diffraction:

A. Improve crystallization conditions

An extensive optimization screening was performed using Detergent Screen, Additive screen, exploiting different crystallization methods (4-20° C temperature, sitting drops). We obtained protein crystals from other conditions without any valuable improvements in diffraction.

Then, we exploited well-established post-crystallization methods to improve crystal quality:

a- Crystal glutaraldehyde cross-linking

Crystal cross-linking increases the robustness of the crystal against mechanical stress, reduces its solubility and can also improve diffraction quality. We treated crystals with glutaraldehyde (vapour diffusion method with a variety % of cross linking agents) with little variations in the diffractive properties of the crystals.

b- Crystal dehydration

Crystallographers have investigated in detail the water-mediated transformations in protein crystals, and it is well appreciated that reduction of solvent content can produce more closely packed and better ordered crystals, extending the resolution of X-ray diffraction patterns (review in Heras & Martin 2005). Crystal dehydration emerges as the treatment that has produced the most remarkable improvements in the diffraction resolution of protein crystals (review in Heras and Martin 2005). Indeed, in case of ERp44 we get small improvement not sufficient to obtain crystals to sufficient resolution.

B. Block the disulfide bond formation in ERp44 after purification

We tried to alkylate free cysteine residues either with NEM (N-ethylmaleimide) or with IAA (iodoacetamide) but after alkylation the protein respectively, aggregated or did not crystallize anymore. We changed the pH in drops from pH 8.5 to 6.5 to reduce the reactivity of cysteine residues, but we obtained small crystals that diffracted poorly. We used other reducing agents such as TCEP and β-mercaptoethanol but we did not improve quality diffraction. We also designed two new mutant versions of ERp44 (Fig. 3.3) lacking of two free cysteine residues:

a) ERp44 SXXS, where the two first reactive cysteines were replaced with serine residues. ERp44 mutant forms aggregates. Despite of our efforts we did not manage to obtain mono-disperse protein to crystallize.

b) ERp44 ΔTRX, where TRX-like a domain containing the two cysteine residues was removed. Although the protein was in monomeric state, extensive crystallization trials failed to obtain high quality crystal.

In all cases we did not manage to improve diffraction-quality. Then we decided to exploit an other approaches to solve ERp44 crystal structure.

C. Remove flexible region from ERp44

The removal of flexible regions from protein is generally used to promote crystallization (Geiger et al., 2008). It is self-evident that molecular flexibility will tend to be detrimental to the formation of a highly ordered crystal lattice. For proteins, such flexibility can arise from large-scale

motions of domains or portions of the polypeptide chain such as surface loops and chain termini. Such motions may be reduced or eliminated by redesigning the protein construct (for instance, by removing disordered terminal extensions). In ERp44 the three thioredoxin-like domains are followed by a disordered C-tail (Fig. 3.2). Based on structural organization of ERp44 and limited proteolysis data (Braakman personal communication), we designed ERp44ΔTail lacking the C-terminal floppy tail that may affects the crystallization (Fig. 3.2 and 3.3). Interestingly, a proteolytic fragment of full-length protein, indicated with a black arrow in Figure 3.8a, co-migrated with construct ERp44ΔTail, indicating that the last 52 residues are susceptible to a protease digestion (Fig. 3.8a) as previously indicated by limited proteolysis data.

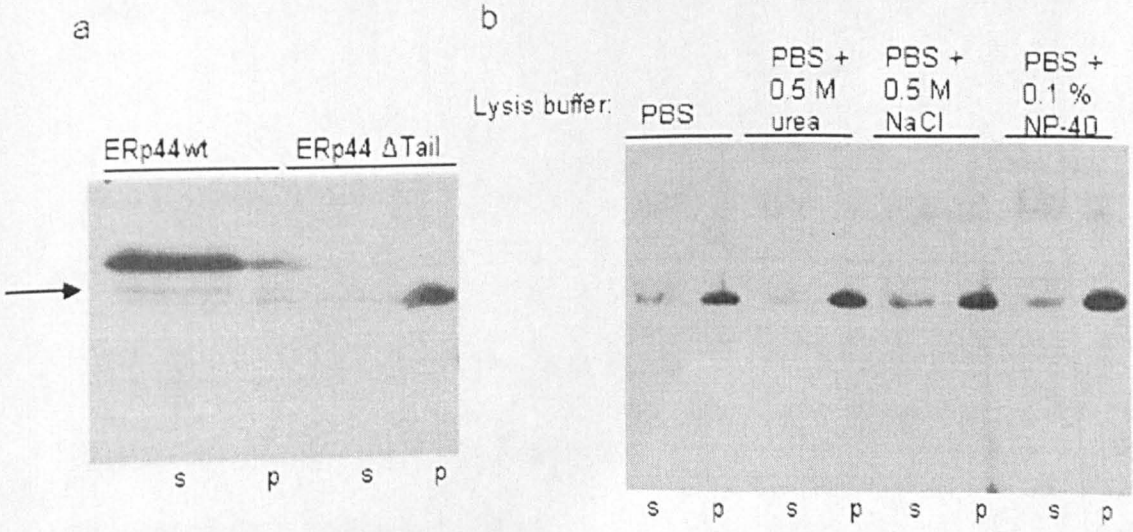


Figure 3.8. ERp44 ΔTail is partially insoluble

A. Protein blot (α-His) of a cell lysate in PBS of BL21 cells expressing ERp44wt or ERp44ΔTail, as indicated. S: soluble fraction; P:insoluble fraction. Arrow indicated proteolytic C-terminal fragment of ERp44 wt.

B. Protein blot (α-His) of a cell lysates with different buffers as indicated, of BL21 cells expressing ERp44ΔTail. S:soluble fraction P. insoluble fraction.

The partial insolubility of ERp44 ΔTail (Fig. 3.8b) strongly suggests that this region, although disordered, is essential for the solubility of the protein at least in *E.coli* (see part 3.2 and 3.3). Despite extensive efforts all these crystals diffracted to 4 Å. During the course of crystallization trials we became aware that the group of Wang in Beijing (China) was working on ERp44 crystal structure. Thus, we started a collaboration with them to join our efforts to solve the structure of human ERp44.

Small motions such as those due to flexible, solvent-exposed amino acid side chains (such as lysine and glutamate residues) can be equally disruptive to a well-ordered crystal lattice (Walter et al., 2006).

The most common approach to reduce surface entropy is the reductive methylation of free amino groups in which primary amines (i.e., lysine residues and the N terminus) are methylated into tertiary amines.

3.2 Crystal structure of human ERp44

3.2.1 Overall structure

The overall structure of ERp44 (Protein data bank accession number 2R2J) resembles a clover composed of three domains (a-, b- and b'-) and a C-tail, which bridges b'- and a-, resulting in a quite compact appearance (Fig. 3.9 a and b). Domains a, b- and b'-shared the TRX-fold. The terminal tail (C-tail, residues 326-377) is highly flexible, ranging from the histidine-rich loop through two short 3_{10} helices: η4 (residues 326-328), and η5 (residues 358-360) filling the cavity formed by α7 and α9 in b'-, short strand β16 (residues 369-370) anti-parallel to β3 in a-, and the ER-retrieval –RDEL motif.

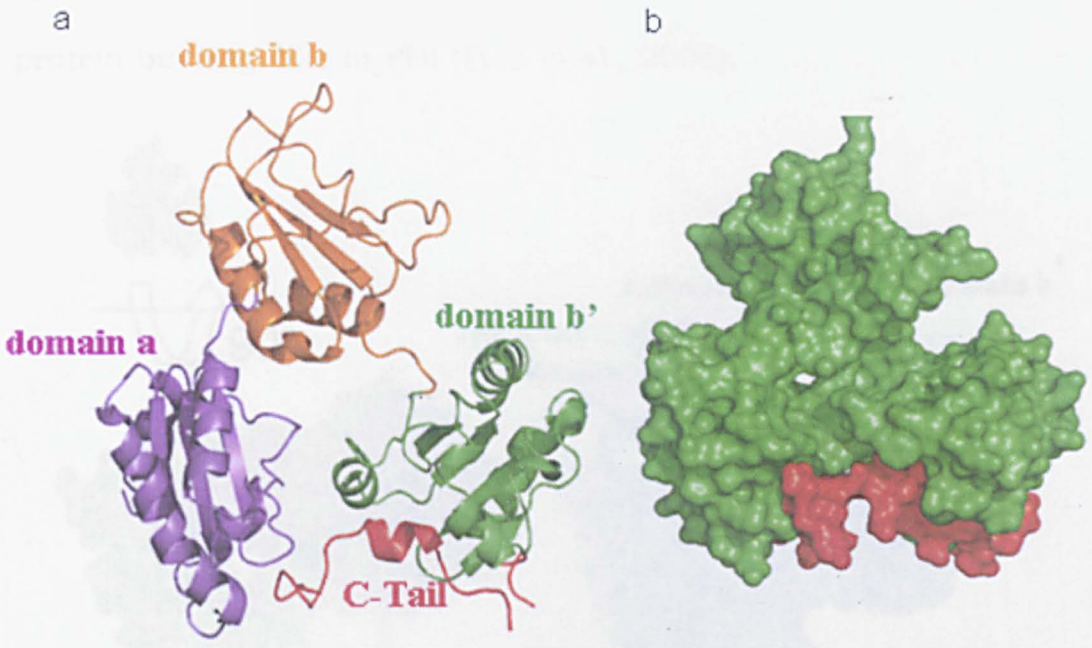


Figure 3.9. ERp44 crystal structure

A. The structure, refined to 2.6 Å resembles a clover with three TRX-like domains (a, b and b'). The C-tail bridges b' and a, resulting in a quite compact appearance. The C-tail (residues 326-377) is flexible, mainly composed of random coils.

B. Surface representation of ERp44. C-tail in red fills the cavity between a- and b- domains. Pictures for surface and cartoon were generated with Pymol.

3.2.2 The -CRFS- motif in domain a

The a domains of ERp44 and yPDI can be superimposed with a root-mean-square deviation (r.m.s.d) of 1.1 Å against 104 aligned Cα atoms of total 109 residues, despite the former having -CRFS- (residues 29-32) instead of the canonical -CXXC- motif. The Cys29 sulphydryl forms hydrogen bonds with the hydroxyl of Ser32 in the equivalent position of Cys64 in yPDI, and the hydroxyl of Thr369 in the C-tail. The -CRFS- motif is surrounded by hydrophobic patches similarly to the one surrounding the -CGHC- motif in yPDI, but is shielded by the C-tail (Fig. 3.10). The

hydrophobic patch around the -CRFS- could also be involved in client protein binding, like in PDI (Tian et al., 2006).

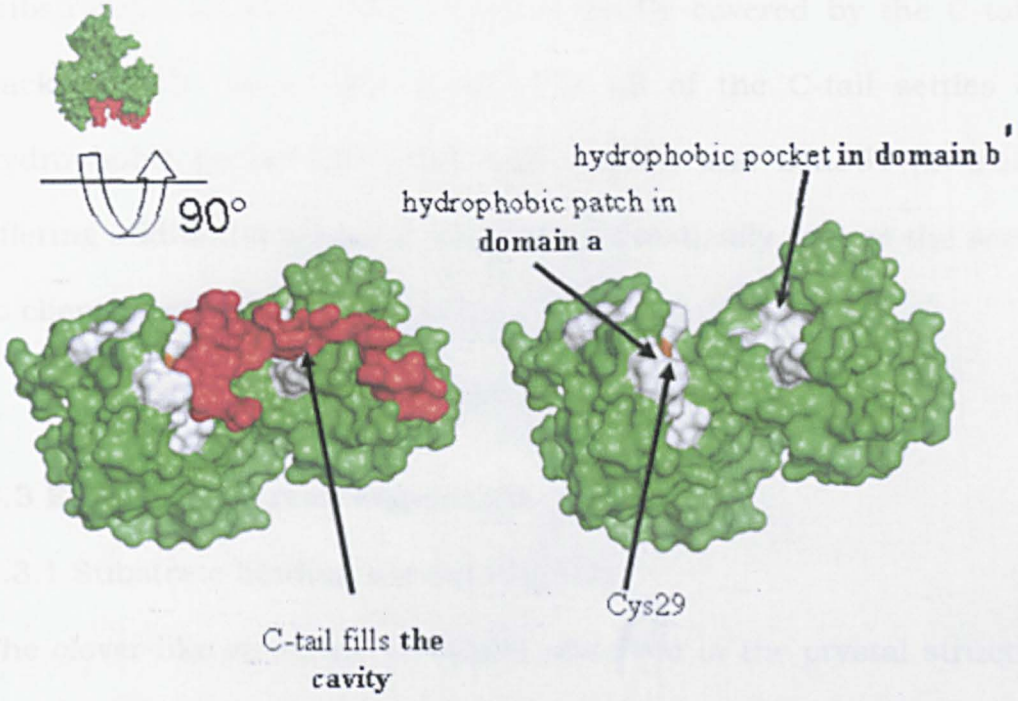


Figure 3.10. C-tail shields Cys29 and hydrophobic pockets

Cys29 (in orange) and hydrophobic pockets (in white) were shielded from the bulk solvent and almost inaccessible for the formation of intermolecular disulfide bonds with client proteins. Pictures for surface were generated with Pymol.

3.2.3 A hydrophobic pocket in b' domain shielded by the C-tail

In both human and yeast PDI, the b' domain is thought to be a primary Substrate Binding Site (SBS) (Klappa et al., 1998) (Tian et al., 2006). A similar hydrophobic pocket in ERp44 is revealed by superposition of the b' domains of ERp44 and yPDI with r.m.s.d of 1.7 Å against 95 aligned Cα atoms out of 110 residues. In ERp44, this pocket is formed by three strands (β12, β13 and β14) at the bottom and two helices (α7 and α9) around with a diameter of ~16 Å, a depth of ~5 Å and a solvent-accessible surface area of ~200 $Å^2$. Sequence alignment based on the structures of hERp44 and yPDI indicated the high similarity of the residues composing

the hydrophobic pocket. The hydrophobic pocket in b'- and the adjacent hydrophobic patch in a- could act as a docking site for ERp44 substrate(s). However, this region is partly covered by the C-tail turned back from b'- to a- (Fig. 3.10). The η3 of the C-tail settles onto the hydrophobic pocket like a lid with Phe358 and Leu361 protruding and offering additional anchors, which may drastically reduce the accessibility to client proteins.

3.3 ERp44 C-tail rearrangements

3.3.1 Substrate binding site accessibility

The clover-like structure of ERp44 observed in the crystal structure most likely represents a non-reactive conformation of the protein (Fig. 3.9 and 3.10). Indeed, Cys29 is shielded from the bulk solvent and almost inaccessible for the formation of intermolecular disulfide bonds with client proteins. The C-tail of ERp44 is actively shielding Cys29 and hydrophobic pockets in domain a and b'- (Fig. 3.10).

To analyze whether the C-tail could have a regulatory role in substrate binding, we engineered a mutated ERp44 lacking the C-tail, hereafter termed ERp44ΔTail (Fig. 3.13).

First of all, we compared ERp44 and ERp44ΔTail for their ability to suppress the aggregation of denatured rhodanese during refolding after dilution (Song & Wang 1995). Since rhodanese has no disulphide bonds, it is possible to analyze the chaperone-like activity of ERp44 or PDI, independently from redox activity. In Figure 3.11, ERp44 wt showed no chaperone-like activity (Fig. 3.11- compare red line to black line).

Conversely ERp44ΔTail (Fig. 3.11- green line) suppressed aggregation as effectively as human PDI (Fig. 3.11- light blue line).

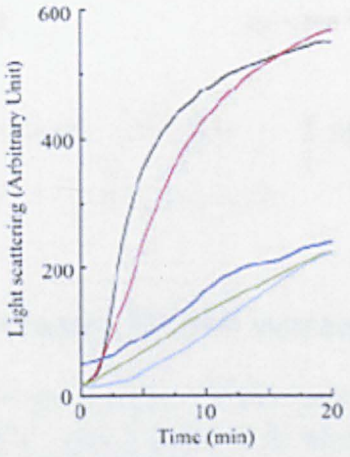


Figure 3.11. ERp44ΔTail showed chaperone-like activity

Time courses of light-scattering changes produced by the aggregation of denatured rhodanese 90 μM during refolding upon dilution (dilution 1:20 in phosphate buffer) in absence (black) and presence of ERp44 wt (red), ERp44ΔTail (green), ERp44T369A (dark blues) or PDI (light blue), as indicated. ERp44 and PDI were added to rhodanese at molar ratio of 8.

Then, we compared the binding ability of ERp44 wt and ERp44ΔTail to interact with Ero1α in a pull-down assay. As showed in Figure 3.12 (compare lanes 1 to 3), ERp44ΔTail bound to Ero1α more efficiently than ERp44 wt. These data indicated that *in vitro* the removal of C-tail enhanced binding and chaperone-like activity of ERp44 suggesting a role of C-tail in regulating Substrate Binding Site (SBS) accessibility.

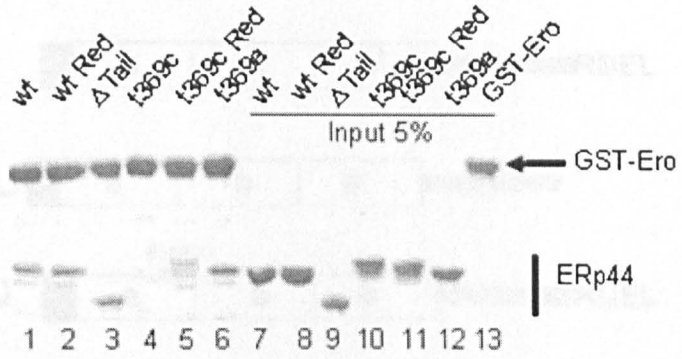


Figure 3.12. Interaction between ERp44 mutants and Ero *in vitro*

Pull-down assay of ERp44 proteins. Glutathiol sepharose 4 fast flow beads were incubated at 25°C, pH 7.5 for 1 h with 10^{-6} M GST-Ero1α and various ERp44 proteins as indicated. Pull-down proteins and input (5% of the total protein used in the pull-down assay) were analyzed by SDS-page under reducing condition. Similar results were obtained in two independent experiments.

To characterize in living cell ERp44 deletion mutant (Fig. 3.13), we transiently transfected different mammalian cell lines (HeLa, CHO, HepG2) with ERp44 wt or ΔTail. The binding activity of ERp44 was then assessed in SDS-page non-reducing conditions and in IP assays. Figure 3.14 shows that ERp44ΔTail displayed an enhanced disulfide-mediated activity towards endogenous proteins respect to wt protein (HMWC- compare lanes 6 to 7).

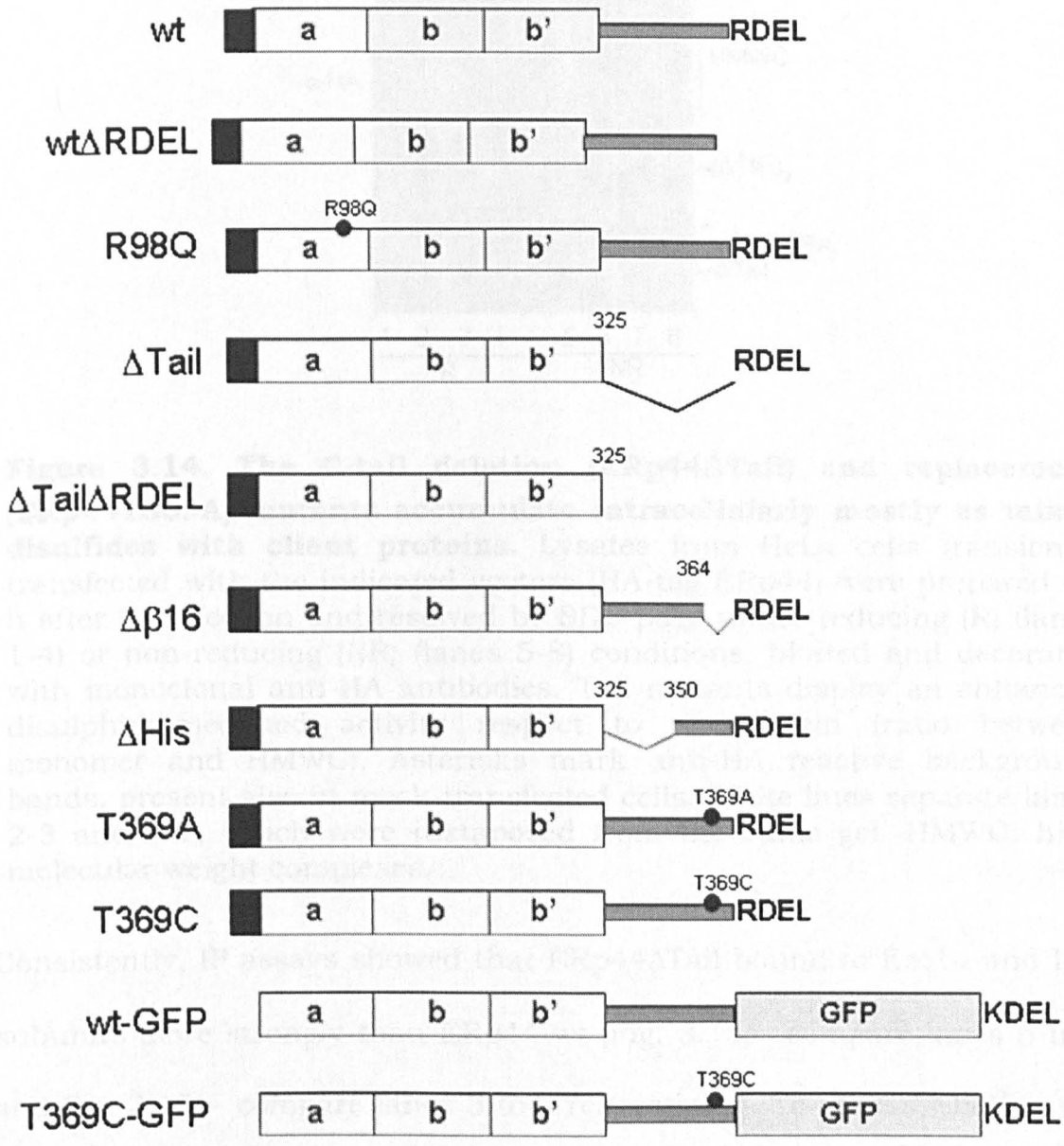


Figure 3.13. ERp44 proteins for mammalian expression

Schematic representation of ERp44 cloned in pCDNA 3.1 vector for mammalian expression. Black square at N-terminus represents HA tag. Black circles are point mutations.

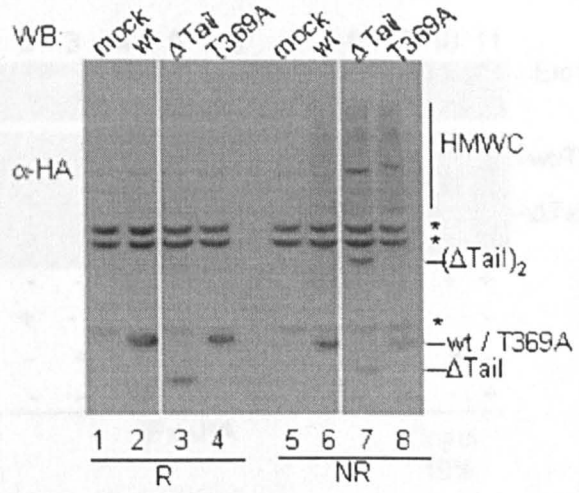


Figure 3.14. The C-tail deletion (ERp44ΔTail) and replacement (ERp44T369A) mutants accumulate intracellularly mostly as mixed disulfides with client proteins. Lysates from HeLa cells transiently transfected with the indicated vectors (HA-tag ERp44) were prepared 48 h after transfection and resolved by SDS-page under reducing (R) (lanes 1-4) or non-reducing (NR) (lanes 5-8) conditions, blotted and decorated with monoclonal anti-HA antibodies. Tail mutants display an enhanced disulphide-mediated activity respect to wt protein (ratio between monomer and HMWC). Asterisks mark anti-HA reactive background bands, present also in mock-transfected cells. White lines separate lanes 2-3 and 6-7, which were juxtaposed from the same gel. HMWC: high molecular weight complexes.

Consistently, IP assays showed that ERp44ΔTail bound to Ero1α and IgM subunits more strongly than ERp44 wt (Fig. 3.15a- compare lanes 6 to 7 and Fig. 3.15b- compare lanes 3 to4, respectively). To assess whether the removal of C-tail enhances the overall binding activity of ERp44 independently from Cys29, we tested the interaction with non-covalent partner ERGIC-53. IP assay in Figure 3.15c showed that ΔTail interacted to ERGIC-53 as much as wt protein. These data indicated that in living cells the removal of the C-tail greatly enhanced the binding properties of ERp44.

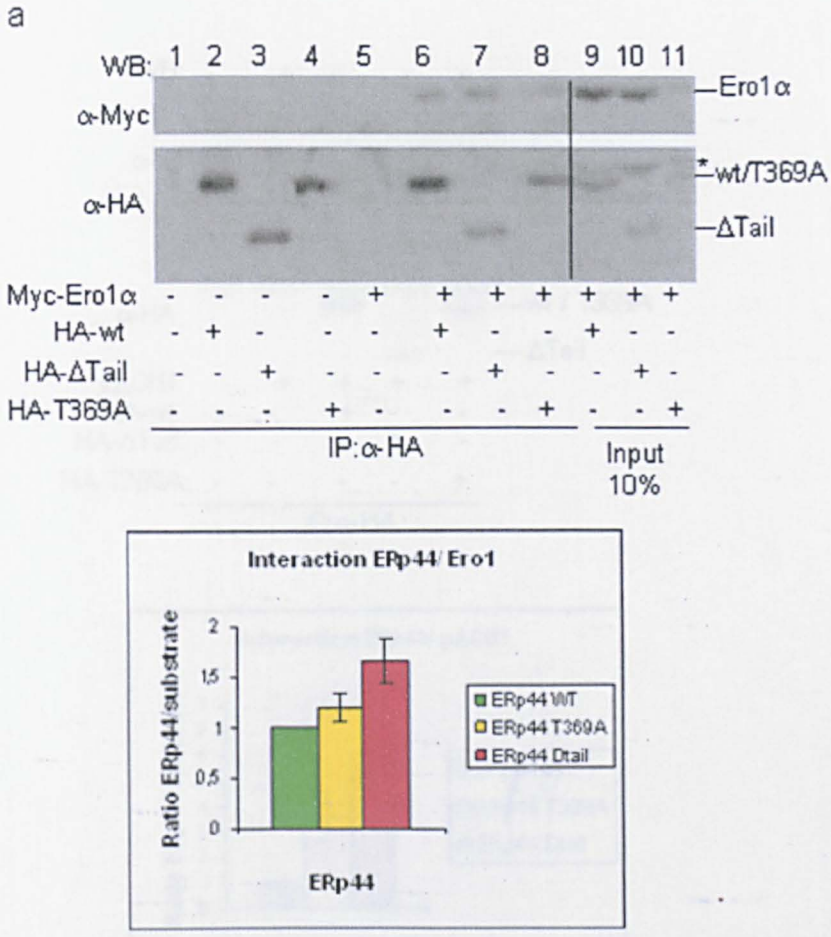


Figure 3.15a. Deletion of the C-tail does not alter ERp44 binding specificities

HeLa cells transfected so as to co-express ERp44, ERp44ΔTail or ERp44T369A and Ero1α (A), μΔCH1 (B) or ERGIC-53 (C) were lysated and immunoprecipitated with anti-HA antibodies. After washing, immunoprecipitates were resolved by SDS-page under reducing conditions, transferred to nitrocellulose and decorated with suitable antibodies to reveal associations (Anelli et al., 2002)(Anelli et al., 2003) (Anelli et al., 2007).

A. The experiment showed is representative of three similar ones. The immunoprecipitates from the total lysates of $3*10^6$ HeLa transfected as indicated, were washed thoroughly and resolved by SDS-page under reducing conditions. Blots were hybridized with anti-Myc to reveal co-immunoprecipitated Ero1α, or with anti-HA antibodies to follow exogenous ERp44. Lanes 9-11 (input) showed an aliquot of the lysates (10^5 cells) before immunoprecipitation, to monitor the expression levels in the different transfectants. Asterisk mark anti-HA reactive background bands. The efficiency of co-immunoprecipitation was determined by densitometric analyses of the Ero1α and ERp44 bands of three independent experiments (see graph below). The graph is densitometric analysis plotting the ratio between ERp44 and Ero1α (normalization: ERp44wt-Ero1α interaction is arbitrarily normalized to 1). The higher is the bar (mean ± SEM), the more effective ERp44 is to interact with Ero1.

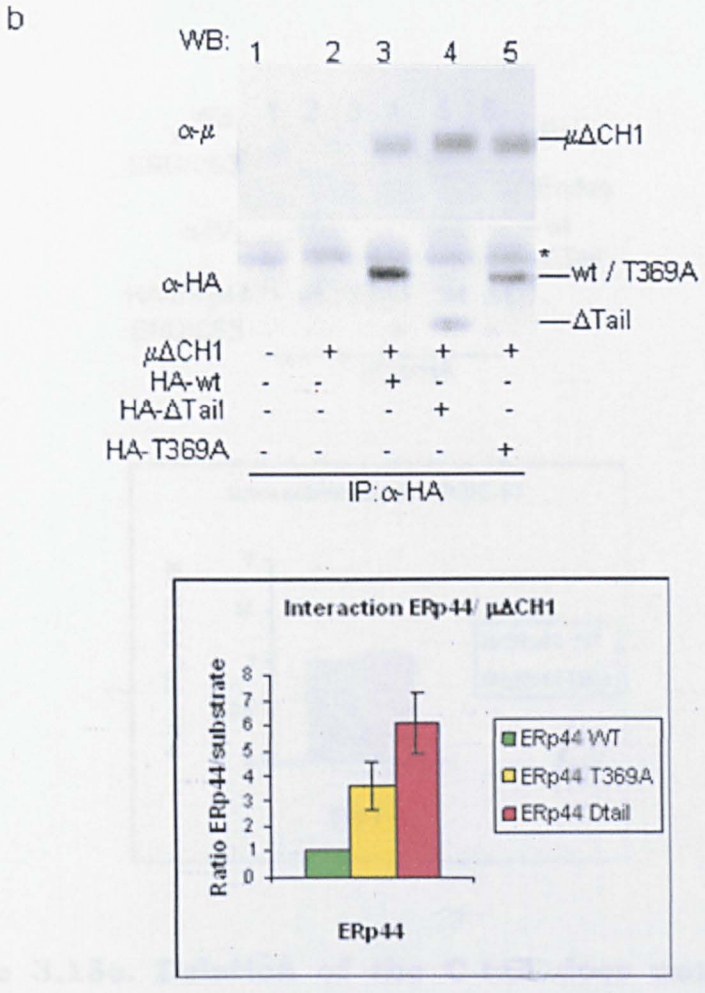
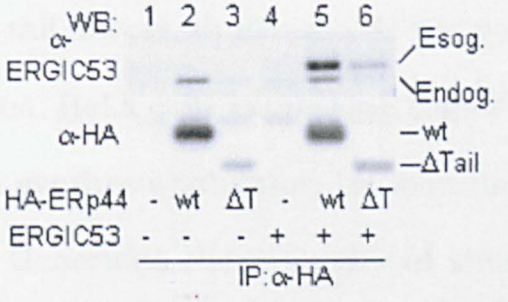


Figure 3.15b. Deletion of the C-tail does not alter ERp44 binding specificities

B. The experiment showed was performed according to a protocol similar to the one described in panel A, using HeLa transfectants expressing ERp44 variants and a known substrate, Ig-μΔCH1 chains (Anelli et al., 2003)(Anelli et al., 2007). Blots were hybridized with anti-μ and anti-HA antibodies. The band indicated with an asterisk consists of the heavy chain of the anti-HA antibody used for immunoprecipitation, that is recognized by the secondary antibody used for detection. The efficiency of co-immunoprecipitation was determined by densitometric analyses of the μΔCH1 and ERp44 bands of three independent experiments (see graph below). The graph is densitometric analysis plotting the ratio between ERp44 and μΔCH1 (normalization: ERp44wt-μΔCH1 interaction is arbitrarily normalized to 1). The higher is the bar (mean ± SEM), the more effective ERp44 is to interact with Ero1α. As described for Ero1α, more μΔCH1 chains are co-precipitated by the tail mutants than by wild-type ERp44.

C



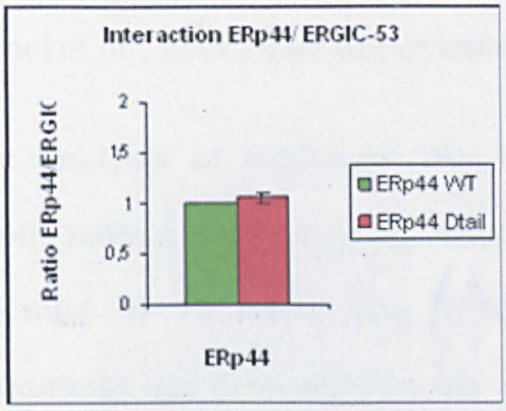


Figure 3.15c. Deletion of the C-tail does not alter ERp44 binding specificities

C. After cross-linking with DSP, a step that increases the efficiency of ERGIC-53/ERp44 co-precipitation (Anelli et al., 2007), the lysates from HeLa cells transfected as indicated were immunoprecipitated with anti-HA, and blots decorated with anti-ERGIC-53 or anti-HA antibodies. Taking into account the different expression levels of wt ERp44 and ERp44ΔTail (ΔT), the results clearly indicate that similar amounts of ERGIC-53 (either endogenous or myc-tagged, glycosylated overexpressed molecules, Anelli et al., 2007) are co-precipitated, in line with the assumption that the association between the two proteins does not involve their substrate-binding domains and Cys29 (Anelli et al., 2007). The graph is densitometric analysis plotting the ratio between ERp44 and ERGIC-53 (normalization: ERp44wt-ERGIC-53 interaction is arbitrarily normalized to 1).

To exclude a possible gain of function non-related to its normal physiology due to the removal of 52 residues from ERp44, we checked whether the removal of C-tail affects, in living cells the stability of the deletion protein. In Figure 3.16a, HeLa cells expressing ERp44 wt or ΔTail were treated with CHX (protein synthesis inhibitor), to monitor the stability of proteins in a time-course. Generally, the intensity of short-lived molecules (i.e. Ik-Ba, orphan Ig subunits) decreased significantly in similar assays (Mancini et al., 2000) (Cenci et al., 2006) (our unpublished results).

Densitometric analysis of SDS-page (Fig. 3.16a) shows that wt and deletion protein have similar stability. Then, we tested the solubility of ERp44 in a mild NP-40 buffer (Fig. 3.16b). Neither ERp44 nor the ERp44ΔTail mutants are detectable in the insoluble fraction when over-expressed in HeLa cells. We can conclude that, although in *E.coli* the protein is partially insoluble (see part 3.1) likely due to hydrophobic pockets accessible to the solvent that could result in aggregation, in living cells it showed similar stability and detergent solubility excluding gross structural alterations (Fig. 3.16).

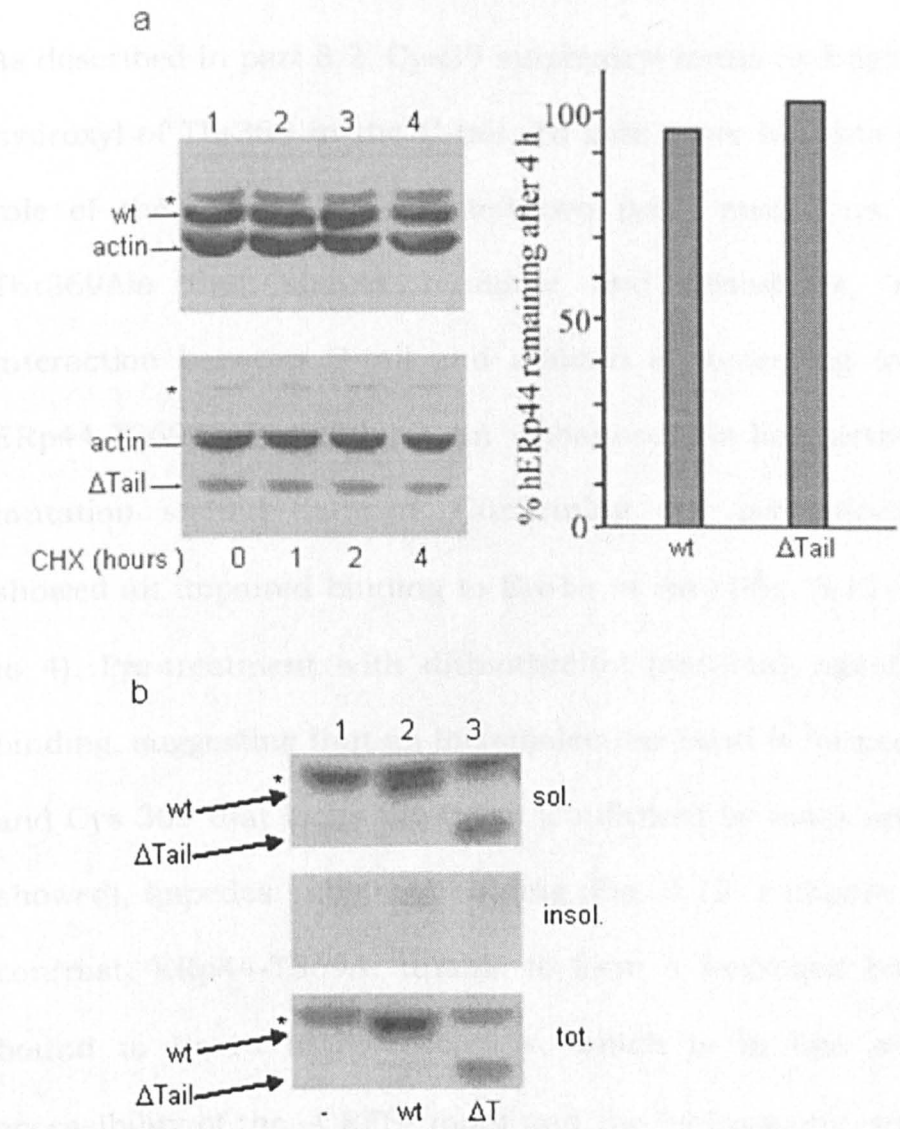


Figure 3.16. Deletion of the C-tail does not cause gross misfolding of ERp44.

A. Stability. HeLa cells expressing ERp44 or ERp44ΔTail were treated with cycloheximide (CHX) for 1, 2 or 4 h as indicated at 48 h after transfection. Aliquots of cell lysates were resolved under reducing conditions in SDS-page, blotted and decorated with monoclonal anti-HA and α-actin antibodies (left panels). The intensity of the ERp44 bands was determined by densitometry and normalized against actin. Asterisks mark anti-HA reactive background bands. The plot in the right panel shows the percentage of ERp44 remaining in cells after 4 h of CHX treatment.

B. Detergent solubility. HeLa cells expressing ERp44 or ERp44ΔTail were lysed in NP-40 (upper two panels) or SDS (lower panel) at 48 h after transfection. NP-40 soluble and insoluble fractions were separated by centrifugation, and aliquots corresponding to the same number of cells resolved by SDS-page, blotted and decorated with anti-HA antibody. Asterisks mark anti-HA reactive background bands.

As described in part 3.2, Cys29 sulphydryl forms hydrogen bonds with the hydroxyl of Thr369 in the C-tail. To gain more insights into a regulatory role of the C-tail we generated two point mutations, Thr369Cys and Thr369Ala that should reinforce and destabilize, respectively, the interaction between C-tail and domain a. According to our predictions ERp44-T369A should have an enhanced binding activity while T369C mutation should dampen. Confirming our predictions, ERp44-T369C showed an impaired binding to Ero1α *in vitro* (Fig. 3.12- compare lanes 1 to 4). Pre-treatment with dithiothreitol (reducing agent) partly restored binding, suggesting that an intramolecular bond is formed between Cys 29 and Cys 369 that locks the C-tail (confirmed by mass spec. analysis; not-showed), impedes substrate binding (Fig. 3.12- compare lanes 2 to 5). By contrast, ERp44-T369A, unable to form a hydrogen bond with Cys 29, bound to Ero1α more efficiently, which is in line with an increased accessibility of the –CRFS- motif and the hydrophobic surfaces (Fig. 3.12- compare lane 1 to 6).

To confirm *in vitro* data, we expressed ERp44-T369A in mammalian cell lines. Replacing Thr369 by alanine increased significantly the fraction of ERp44 that forms mixed disulphides with endogenous proteins in HeLa cells (Fig. 3.14–compare lanes 6 to 8) respect to wt protein. Moreover, IP assay (Fig. 3.15) showed that ERp44-T369A interacted with Ero1α and with IgM subunits (Fig. 3.15a and b, respectively) more strongly than wt protein.

This data demonstrated the interactions between C-tail and domain a (ERp44 T369A-destabilization; ERp44 T369C strengthening) could

modulate the binding properties of ERp44 suggesting a role of C-tail in regulating SBS accessibility during protein quality control in the Early Secretory Compartments (ESC).

We investigated the mechanism allowing ERp44 to expose its binding sites for its physiological function by means of molecular dynamics (MD) simulations, a well suited approach to appreciate the dynamic behaviour of macromolecules (Karplus & Kuriyan 2005) (Rueda et al., 2007). Monitoring the conformations of ERp44 under constant pressure and temperature for 10 ns (sampling of the NPT ensemble) for the wild type ERp44 protein revealed a stable, compact clover structure (Fig. 3.17- wt). A principal component analysis of the trajectory highlighted a propensity of the individual TRX-like domains to undergo rigid-body rotational movements. The C-tail of ERp44 apparently acts as a "molecular glue", forcing a- and b'- domains to undergo correlated motions. The interactions between strands β3 and β16, the hydrogen bonds between Arg367 and residues of the β5-α4 loop, and the side chain interactions between Thr369 and Cys29 do not allow an "opening" of the C-tail (at least in the used time scale). Hence, we explored whether the removal of the last 9 residues of ERp44 (including β16 strand) could destabilize the interface, and facilitate the opening of the ERp44 C-tail (hereafter the mutant is named ERp44Δβ16). Indeed, after 10 ns of NPT MD simulation the ERp44Δβ16 mutant undergoes a significant structural rearrangement (Fig. 3.17- Δβ16). Domain a partially reorients independently of b'-, and helix α2 (in purple) bearing the active site residue Cys29 assumes a regular, straightened conformation to fully expose the thiol to the solvent (Fig.

3.17). This conformation is suggestive of a highly reactive sulphydryl, and may thus represent an active conformation of ERp44.

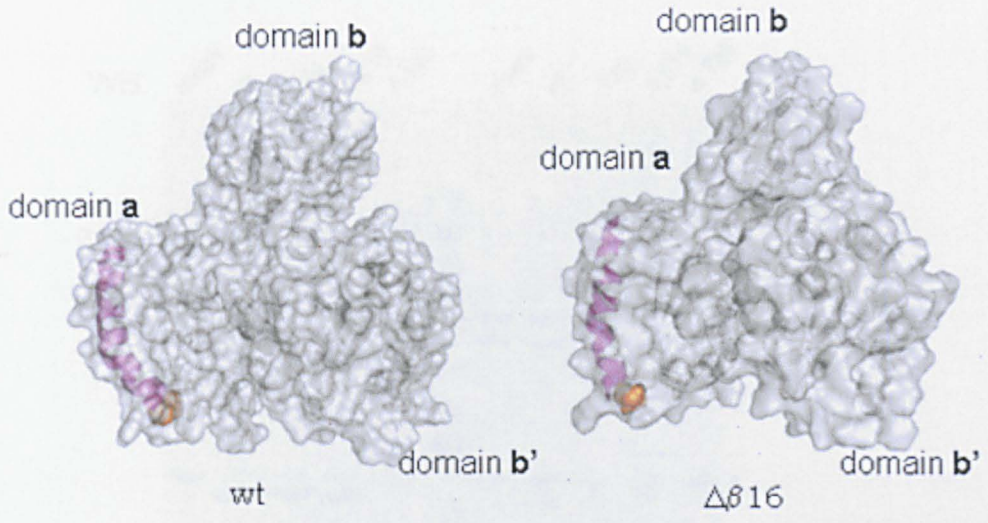


Figure 3.17. ERp44Δβ16 mutant undergoes a significant structural rearrangement

Final structures of ERp44 wt and ERp44 Δβ16 after 10 ns of NPT MD simulation. The secondary structure cartoons are included inside transparent surface representations of ERp44 wt and ERp44Δβ16 mutant. The latter undergoes a significant structural rearrangement respect to wild type protein: domain a partially reorients independently of b', and helix α2 (in pink) bearing the active site residue Cys29 (orange sphere) assumes a regular, straightened conformation to fully exposes the thiol to the solvent. Pictures for cartoons and surfaces, were generated with Pymol.

We tested the effect of the expression of ERp44Δβ16 mutant (-RDEL motif is maintained) in HeLa cells. Compared to the wild type protein (Fig. 3.18- compare lanes 7 to 9), in SDS-page non-reducing conditions this mutant displays a substantially higher tendency to intermolecular disulfide bond formation, including homodimerization yielding a ΔTail-like phenotype (Fig. 3.18- compare lanes 9 to 10). The removal of the last 9 residues of ERp44 enhances the disulfide-mediated activity similarly to the full removal of the C-tail, as expected from the conformations sampled in the

MD simulation. Moreover, the results hint at a primary role of the domain a as a primary binding site for client proteins.

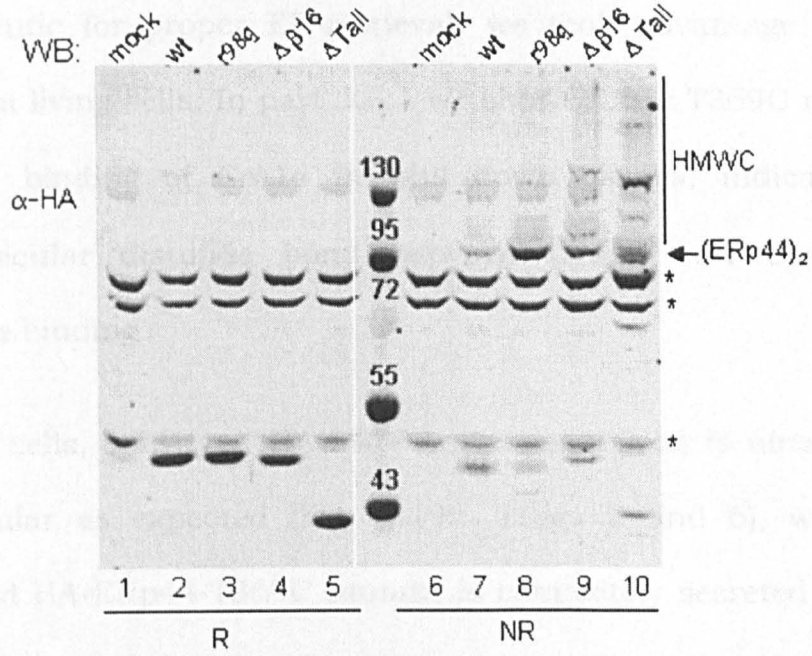


Figure 3.18. Removal of the last 9 residues of ERp44 enhances the disulfide-mediated activity towards endogenous client proteins in living HeLa cells. Lysates from HeLa cells, transiently transfected with HA-ERp44 wt, HA-ERp44Δβ16, HA-ERp44ΔTail, ERp44R98Q or empty vector as indicated, were resolved by SDS-page reducing (R) non-reducing conditions (NR), blotted and decorated with monoclonal anti HA antibodies. Tail mutants display an enhanced disulphide-mediated activity respect to wt protein (ratio between monomer and HMWC). R98Q mutant accumulates as homodimers. The lane between 6 and 7 represents the molecular weight marker. Asterisks mark anti-HA reactive background bands, present also in mock-transfected cells. HMWC: High Molecular Weight Complexes.

3.3.2 -RDEL motif availability in ERp44

Most soluble ER proteins (chaperones and folding enzymes) contain ER-retrieval motifs at the C-terminus that are recognized by KDEL-Receptor (KDEL-R) implying that the motif should be exposed to the solvent, and thus accessible for the interaction with its receptor. Indeed, the -KDEL motif is not visible in the crystal structure of ER resident protein presumably due to high flexibility (ERp44 in Fig. 3.9) (yPDI-Tian et al., 2006) (ERp29-Barak et al., 2009). Nevertheless, in the crystal structure of

ERp44 the "closed" conformation of the C-tail likely positions –RDEL motif very close to domain a and so may partially hinders its availability to KDEL-R. In order to investigate whether the C-tail rearrangements are propaedeutic for proper ER-retrieval, we took advantage of Thr369Cys mutant in living cells. In part 3.3.1 we showed that T369C mutant has an impaired binding of Ero1α in pull down assays, indicating that the intramolecular disulfide bond between Cys29 and Cys369 impedes substrate binding.

In HeLa cells, wild type HA-ERp44 over-expressed is almost exclusively intracellular as expected (Fig. 3.19a- lanes 2 and 6), while the over-expressed HA-ERp44-T369C mutant is completely secreted in the culture medium, to levels comparable with overexpressed HA-ERp44ΔRDEL that lacks the ER retrieval signal (Fig. 3.19a- compare lanes 7 to 8). However, endogenous ERp44 is fully retained intracellularly (faster migrating bands in lanes 1-4) indicating that KDEL-R is not saturated by the over-expression of HA-ERp44. These data suggest that the "closed" conformation of the C-tail may hamper KDEL-R recognition of the ERp44 C-terminal motif. Alternatively, it may suggest a requirement for interaction of ERp44 with its substrates for its retrieval in the ER.

In accordance to the first hypothesis, a chimerical molecule ERp44-T369C-GFP-KDEL (Fig. 3.13), in which a GFP has been placed between the end of the C-tail and the -KDEL sequence (providing a sort of "spacer" for fully expose –KDEL motif), is partially intracellularly retained (Fig. 3.19b-lanes 3 and 5).

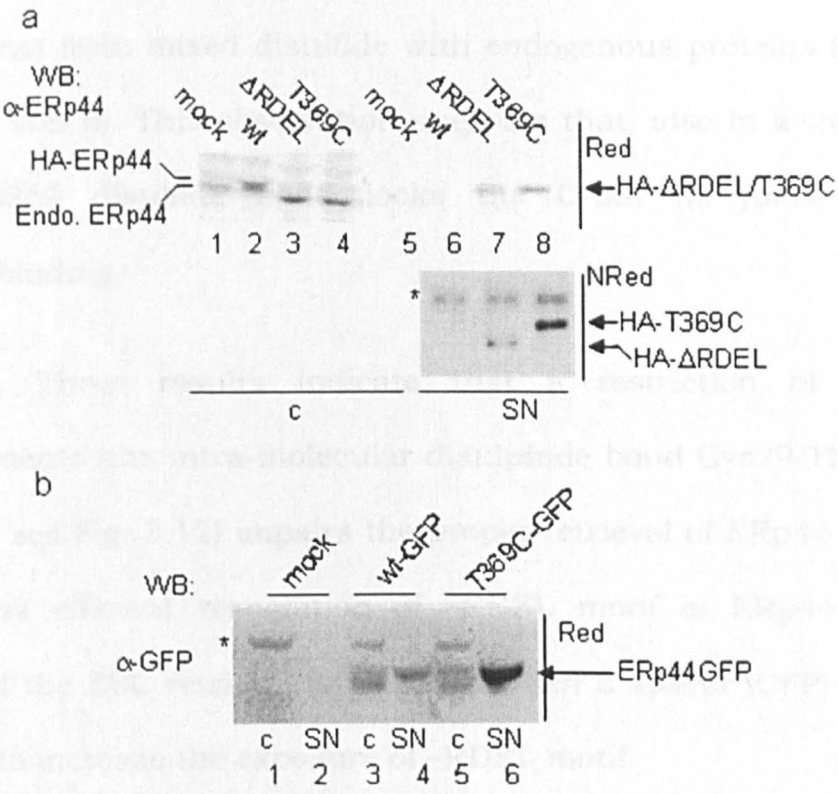


Figure 3.19a and b. ERp44 T369C is fully secreted in living HeLa cells.

A. Lysates (c) from HeLa cells, transiently transfected with HA-ERp44 (wt), HA-ERp44ΔRDEL (ΔRDEL), HA-ERp44-T369C (T369C) or empty vector as indicated were resolved by SDS-page under reducing (Red) or non-reducing (NRed), conditions, blotted and decorated with monoclonal anti ERp44. The culture supernants (SN) of 16h were precipitated with TCA and then resolved under reducing condition (Red) or IP with α-HA antibodies and then resolved under non-reducing conditions (NRed), as indicated. The band indicated with an asterisk consists of the heavy chain of the anti-HA antibody used for immunoprecipitation, that is recognized by the secondary antibody used for detection.

B. Lysates (c) from HeLa cells, transiently transfected with ERp44-GFP-KDEL (wt-GFP), ERp44-T369C-GFP-KDEL (T369C-GFP) or empty vector as indicated were resolved by SDS-page under reducing (Red)conditions, blotted and decorated with monoclonal anti GFP. The culture supernants (SN) were precipitated with TCA and then resolved under reducing condition (Red). Asterisks mark anti-GFP reactive background bands, present also in mock-transfected cells. The doublet present in ERp44-GFP could represent a post-translational modification that occurs in the chimera.

The SDS-page non-reducing conditions in Figure 3.19c shows that intracellular ERp44-GFP-KDEL displays a disulphide-mediated activity

with endogenous proteins while intracellular ERp44-T369C-GFP-KDEL seems do not form mixed disulfide with endogenous proteins (see HMWC in lanes 5 and 6). This observation suggests that, also in living cells, the Cys29-Cys369 disulfide bond locks the C-tail *de facto* hampering substrate binding.

Altogether, These results indicate that a restriction of the C-tail rearrangements (the intra-molecular disulphide bond Cys29-Thr369 locks the C-tail- see Fig. 3.12) impairs the proper retrieval of ERp44 most likely due to less efficient recognition of –RDEL motif of ERp44. A partial recovery of the ESC retrieval is observed when a spacer (GFP) is provided to ERp44 to increase the exposure of –RDEL motif

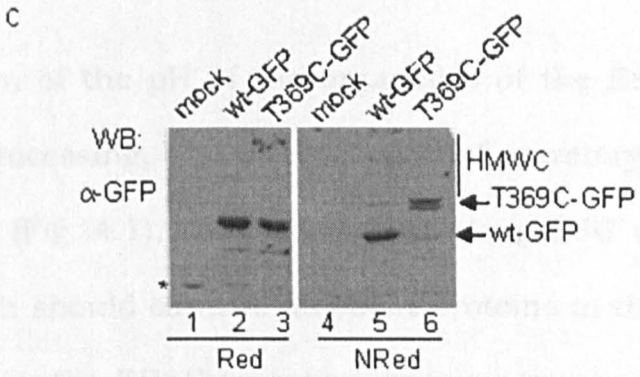


Figure 3.19c. ERp44 T369C is fully secreted in living HeLa cells. Lysates from HeLa cells, transiently transfected with ERp44-GFP-KDEL (wt-GFP-KDEL), ERp44-T369C-GFP-KDEL (T369C-GFP) or empty vector as indicated were resolved by SDS-page under reducing (Red)or non-reducing conditions (NRed), blotted and decorated with monoclonal anti GFP. The limited shift of molecular weight of T369C-GFP in non-reducing SDS-page respect to wt-GFP in Fig. 3.19c (compare lanes 5 to 6) is suggestive of Cys29-Cys369 bond rather than being a covalent mixed disulphide. The same shift is present in HA-ERp44-T369C in Fig. 3.19a in comparison with HA-ERp44 (non-reducing panel lanes 7 to 8). Faster migrating bands may represent protein degradation of GFP chimeras. Asterisks mark background bands. HMWC: High Molecular Weight Complexes.

.

3.4 ERp44 regulation in living cells

3.4.1 ESC pH gradient regulates ERp44 activity

To investigate the mechanisms regulate that C-tail rearrangements, we focused our attention on the pH gradient along the ESC. Indeed, the luminal pH is not homogeneous throughout the secretory pathway (Paroutis et al., 2004) whereas the pH of the endoplasmic reticulum is near neutral, similar to that of the cytosol, downstream compartments become progressively more acidic. The *cis*-Golgi complex is significantly more acidic (pH ~ 6.7) than the ER and the acidification becomes more marked in subsequent cisternae of the Golgi complex, reaching pH ~ 6.0 in the *trans*-Golgi network (Paroutis et al., 2004).

The regulation of the pH of the organelles of the ESC is critical for the trafficking, processing, and glycosylation of secretory proteins. According to our model (Fig. 4.1), the pH gradient along ESC might control ERp44 activity, which should capture its client proteins in the Golgi complex and release them in the ER. Supporting evidence comes from the observation that ERp44-IP_3R1 interaction is pH−, redox state−, and ER Ca^{2+}-dependent (Higo et al., 2005).

Ammonium chloride treatment is widely exploited to raise the ESC luminal pH in living cells (Axelsson et al., 2001) (Appenzeller-Herzog et al., 2004) (Maeda et al., 2008). According to our model (Fig. 4.1), a neutralisation of the pH gradient along the ESC should dampen thiol-mediated retention of ERp44.

Our data indicate that treatment with ammonium chloride results in an increased formation of homodimers and covalent complexes with natural substrates in overexpressed wt protein compared to physiological conditions (Fig.3.20- compare lanes 6 to 8).

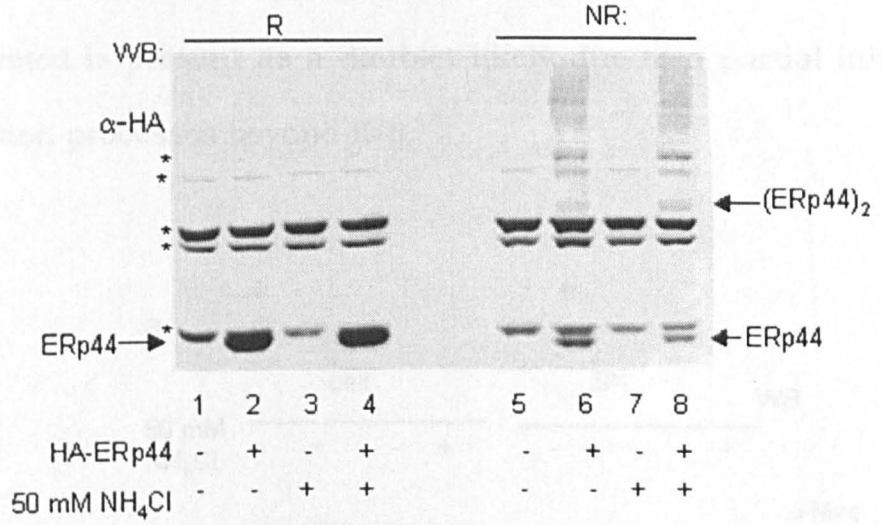


Figure 3.20. Ammonium chloride treatment induces ERp44 homodimerization.

Lysates from HeLa cells, transiently transfected with HA-ERp44 wt or with an empty vector as indicated, were resolved by SDS-page under reducing (R) non-reducing (NR) conditions, blotted and decorated with anti-HA antibodies. The arrows indicate ERp44 homodimer, which is induced after NH_4Cl treatment (ratio dimer/monomer increases upon NH_4Cl treatment). Similar results were obtained in 3 independent experiments. Asterisks mark anti-HA reactive background bands, present also in mock-transfected cells.

The pH sensitivity of ERp44 activity was assessed in functional assays exploiting our well-established assay for TMR activity in which overexpressed ERp44 retains overexpressed Ero1α (Anelli et al., 2003)(Otsu et al., 2006).

As showed in Figure 3.21a (and in the densitometric analysis graph in Fig. 3.22), ERp44-mediated Ero1α retention is weakened by NH_4Cl treatments in HeLa cell lines (Fig. 3.21a- compare lanes 7 to 8) (confirmed also in

CHO cell lines; not-showed). Thus, more Ero1α accumulates extracellularly, whilst the secretion of α1AT (Fig. 3.21c), a protein that does not bind ERp44 (not-showed), is inhibited by NH_4Cl (Fig. 3.21c-compare lanes 7 to 8) (a condition known to perturb glycan processing and secretion reported in Axelson et al., 2001 and in Maeda et al., 2008; Ero1 secreted is present as a doublet likely due to a partial inhibition of glycosylation processes beyond ER).

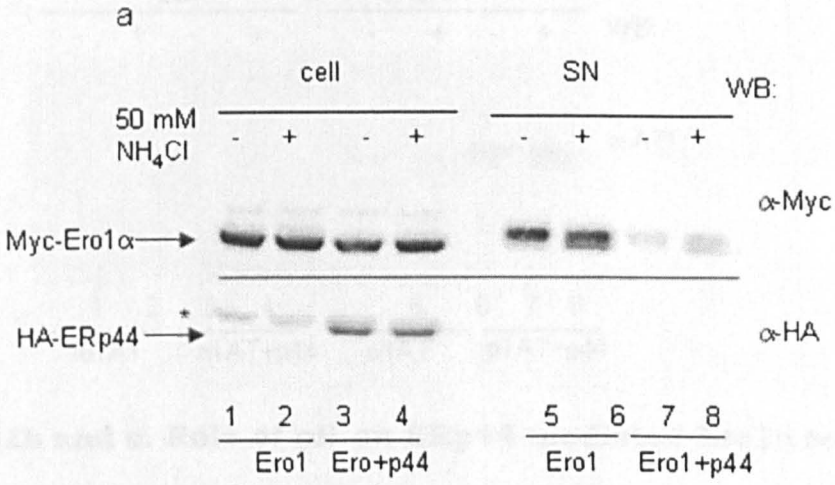


Figure 3.21a. Role of pH on ERp44-mediated Ero1α retention

A. As predicted by our model NH_4Cl treatment makes ERp44 less active in retaining Ero1α. Supernatants after 4 h of culture in Optimem (SN) and cell lysates (cell) from HeLa cells transfected with Myc-Ero1 alone or Myc-Ero1 + HA-ERp44 wt (p44) as indicated, were analyzed by Western blotting with monoclonal antibodies against Myc (Ero1α) or HA (ERp44). When over-expressed, Ero1α is partially secreted (lane 5). The over-expression of ERp44 drastically reduced secretion (compare lanes 5 with 7). ERp44-mediated retention is weakened by NH_4Cl treatment (compare lanes 7 with 8). The appearance of secreted Ero1α bands with different mobility reflects altered glycan processing in NH_4Cl-treated cells (Maeda et al., 2008). See Fig. 3.22 for the densitometric quantification of 10-22 experiments performed like the one showed in SDS-page, plotting the ratio between intracellular (IN) and secreted (OUT) Ero1α.

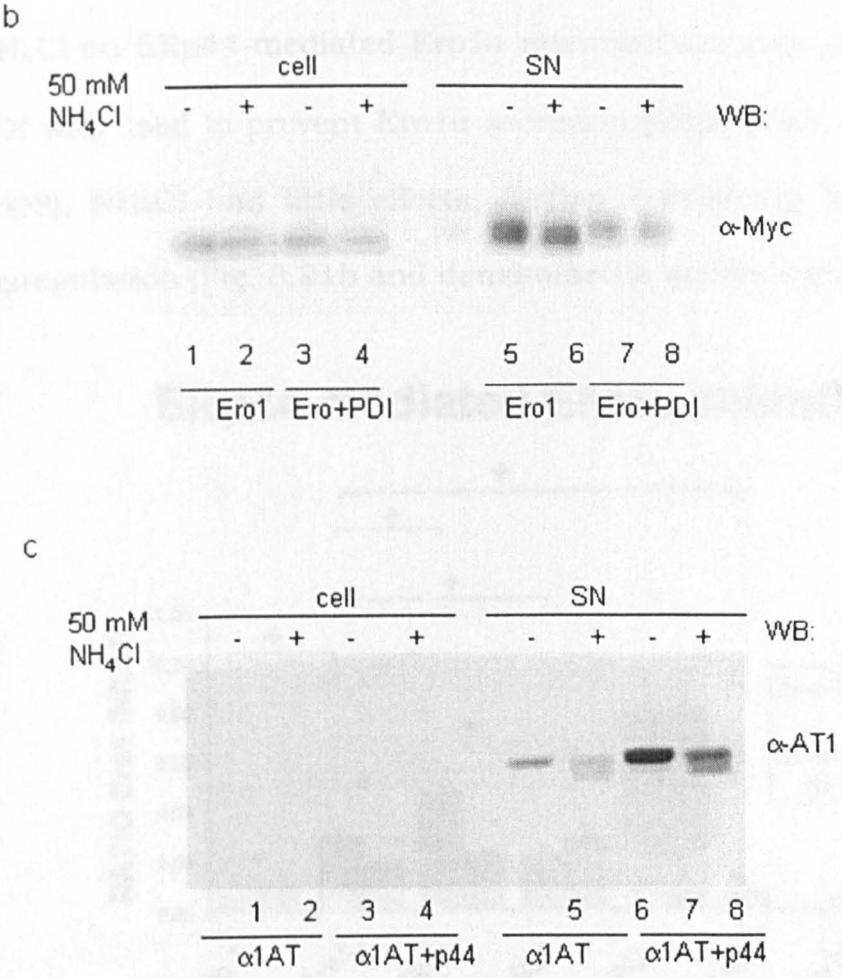


Figure 3.21b and c. Role of pH on ERp44-mediated Ero1α retention

B. Supernatants (after 4 h of culture in Optimem) and cell lysates from HeLa transfected with Ero1α-Myc alone or Ero1α-Myc + PDI as indicated, were analyzed by Western blotting with monoclonal antibodies against Myc (Ero1α).

The over-expression of PDI drastically reduced secretion (compare lanes 5 with 7). PDI-mediated retention is little affected by NH_4Cl treatment respect to ERp44 (compare lanes 7 with 8). See Fig. 3.22 for densitometric quantification of 5 experiments performed like the one showed in SDS-page, plotting the ratio between intracellular (IN) and secreted (OUT) Ero1α. The higher is the bar, the more effective PDI is to retain Ero1α .

C. Supernatants (after 4 h of culture in Optimem) and cell lysates from HeLa cells transfected with α1AT or α1AT + HA-ERp44 as indicated, were analyzed by SDS-page, blotted and decorated with anti αAT1 monoclonal antibody. When over-expressed, α1AT, a protein which is not known to be an interactor of ERp44, is fully secreted (lane 5). The secretion of α1AT is inhibited by NH_4Cl (compares lane 5 with lane 6 and 7 with 8). Considering that NH_4Cl inhibits secretion of α1AT and Ero1 expressed in the same cells, owing to a general effect on processing and secretion (Maeda et al., 2008), the effects of pH on ERp44-mediated Ero1α secretion are more marked.

Considering a less efficient secretory pathway, the inhibitory effects of NH_4Cl on ERp44-mediated Ero1α retention are even more marked. When PDI was used to prevent Ero1α secretion (Otsu et al., 2006) (Fraldi et al., 2008), NH_4Cl had little effects, further confirming the specificity of pH disregulation (Fig. 3.21b and densitometric analysis graph in Fig. 3.22).

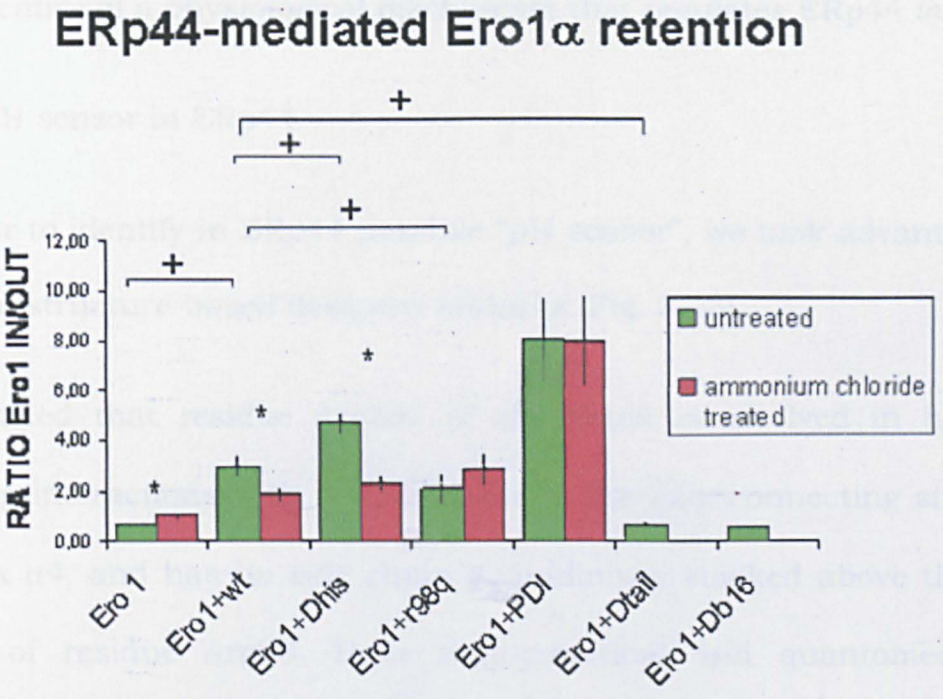


Figure 3.22 ERp44-mediated Ero1α retention

Plot is densitometric analysis of Fig. 3.21a and b and Fig. 3.24 of 5-22 experiments performed like the ones showed in SDS-page in the figures, plotting the ratio between intracellular (IN) and secreted (OUT) Ero1α. The higher is the bar (mean ± SEM), the more effective ERp44 is to retain Ero1α. Asterisk: two tailed T-test, *p<0.05. Plus: one-way ANOVA with Bonferroni's post tests, +p<0.01. Differences were considered significant at $p < 0.05$.

Fig. 3.21:NH_4Cl treatment weakens the TMR activity of ERp44 but has little effect on PDI.

Fig. 3.25: ERp44ΔHis displayed an enhanced TMR activity respect to wt protein, while ERp44ΔTail and ERp44Δβ16 were partially inhibited.

NH_4Cl treatment does not grossly affect the localization of overexpressed HA-ERp44, which accumulated primarily in ER (IF in Fig. 3.26a). Moreover, NH_4Cl treatment seems that does not affect KDEL-R activity since ERp44 was not detected in the medium (non-showed).

Given the known pH differences in the ESC organelles, these results may be indicative of a physiological mechanism that regulates ERp44 *in vivo*.

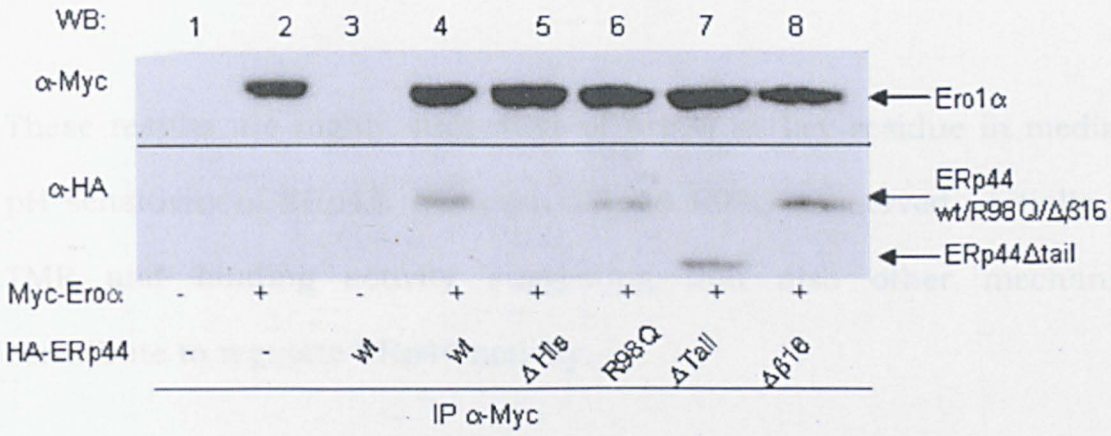
3.4.2 pH sensor in ERp44

In order to identify in ERp44 possible "pH sensor", we took advantage of a panel of structure-based designed mutants (Fig. 3.13).

We noticed that residue Arg367 of the C-tail is involved in hydrogen bonding interactions with the backbone of the loop connecting strand β5 to helix α4, and has its side chain guanidinium stacked above the same group of residue Arg98. Both semi-empirical and quantomechanical analyses suggested a shift of the pKa of Arg98 towards lower values, making this residue pair a candidate as a pH sensor in the different compartments of the ESC.

To test the role of these residues in pH mediated activity of ERp44 we replaced Arg98 with a Gln (mimicking the deprotonated state) yielding ERp44 R98Q. In agreement with our prediction, Arg98Gln mutation seemed to partially dampen TMR activity of ERp44 (Fig. 3.25- compare lanes 10 to 13). Owing to partial inhibition of this activity, ERp44 R98Q is less affected by NH_4Cl treatment respect to wt protein (densitometric analysis Fig. 3.22). The partial inhibition of TMR activity is also confirmed

by IP assay in which ERp44 R98Q interacted to less extent to Ero1α compared to wt protein (Fig 3.23- compare lanes 4 to 6)



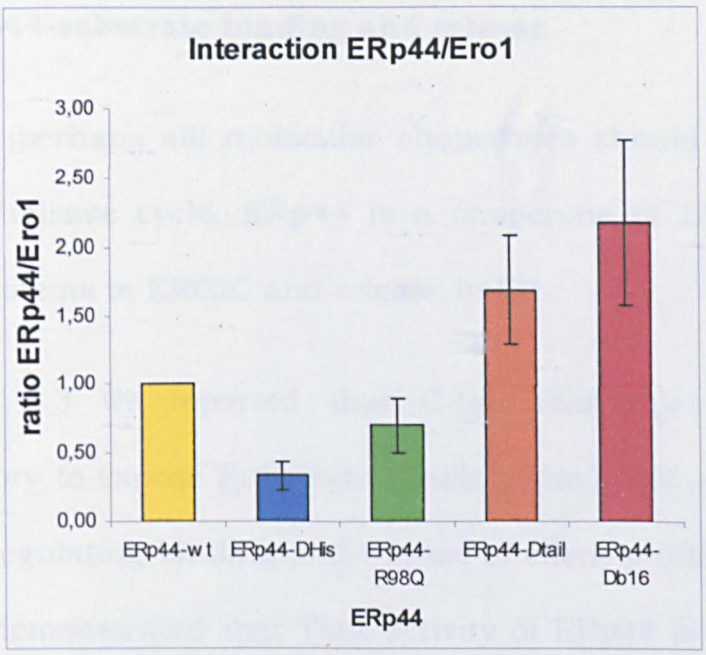


Figure 3.23. Interaction between Ero1α and ERp44 mutants

HeLa cells were transfected with the indicated plasmids. Cell lysates were immunoprecipitated (IP) with α-Myc sepharose beads, and then resolved by SDS-page, blotted and decorated with anti-HA (ERp44) and anti-Myc (Ero1α) antibodies, as indicated. The efficiency of co-immunoprecipitation was determined by densitometric analyses of the Ero1α and ERp44 bands of four independent experiments (see graph below). The graph is densitometric analysis plotting the ratio between ERp44 and Ero1α (normalization: ERp44wt-Ero1α interaction is arbitrarily normalized to 1). The higher is the bar (mean ± SEM), the more effective ERp44 is to interact with Ero1α. Note that the tail ERp44 mutants interact more strongly with Ero1α than wild-type ERp44 does whereas ΔHis and R98Q mutants interact less with Ero1 respect to wt protein.

Peculiarly, in non-reducing SDS-page (Fig. 3.18- lanes 12) ERp44 R98Q accumulates mainly in covalent homodimeric species, enriched in NH_4Cl treated cell expressing ERp44 wt (Fig. 3.20).

These results are highly suggestive of Arg98 as key residue in mediating pH sensitivity of ERp44. However, ERp44 R98Q conserved partially both TMR and binding activity suggesting that also other mechanisms contribute to regulate ERp44 activity.

3.5 ERp44-substrate binding and release

Most of (perhaps all) molecular chaperones should undergo a substrate binding/release cycle. ERp44 is a chaperone of ESC that should bind client proteins in ERGIC and release in ER.

In part 3.3 we reported that C-tail rearrangements of ERp44 are mandatory to expose Substrate Binding Site (SBS) and thus could play a role in regulating binding and release of client proteins. Moreover, in part 3.4 we demonstrated that TMR activity of ERp44 is regulated by ESC pH gradient. To further analyze the role of C-tail in chaperone properties of ERp44 we took advantage of a panel of structure-based mutants.

Moreover, we showed that ERp44Δβ16 displayed an enhanced disulfide-mediated activity of ERp44 (Fig. 3.18). Consistently, Figure 3.23 shows that in IP assay ERp44Δβ16 displayed an enhanced binding activity towards Ero1α in HeLa cells respect to wt protein (Fig. 3.23- compare lanes 4 to 8 and its densitometric analysis below).

In order to identify other "hot spot" for ERp44 binding/release cycle, we removed a conserved His loop (yielding ERp44ΔHis- Fig. 3.13) in the C-tail, invisible in the crystal structure likely due to its flexibility (see part 3.2 and Fig. 3.9). Often disordered regions in protein serve as docking site for ligand partners, and undergo a structural rearrangement upon complex formation (Luan et al., 2000) (Fink 2005)(Dierks et al., 2005).

We tested the effect of the expression of ERp44ΔHis in HeLa cells. Compared to the wild type protein, ERp44ΔHis displayed clearly lower tendency to intermolecular disulfide bond formation towards endogenous protein (Fig. 3.24).

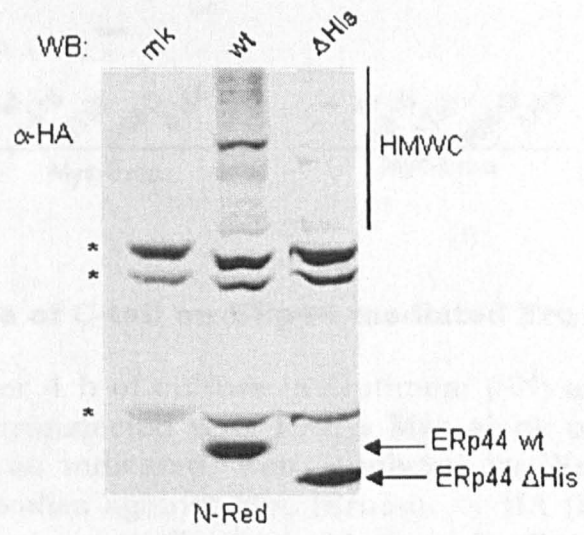


Figure 3.24. The removal of the His-loop diminishes the disulfide-mediated activity of ERp44-

Lysates from HeLa cells transiently transfected with HA-ERp44 wt, ERp44ΔHis or empty vector were resolved by non-reducing conditions (NRed), blotted and decorated with monoclonal anti HA antibodies. Note that ERp44ΔHis displays a reduces disulphide-mediated activity respect to wt protein (ratio between monomer and HMWC). Asterisks mark anti-HA reactive background bands, present also in mock-transfected cells. HMWC:High Molecular Weight Complexes.

Consistently, the removal of the His loop, although remote from SBS in domain a, dampens the binding activity of ERp44 to Ero molecules (IP assay; Fig. 3.23– compare lanes 4 to 5 and densitometric analysis below).

These data suggested that C-tail contributes to regulating the binding properties of ERp44 where the removal of β16 and His-loop make the protein, hyper-reactive and hypo-reactive, respectively.

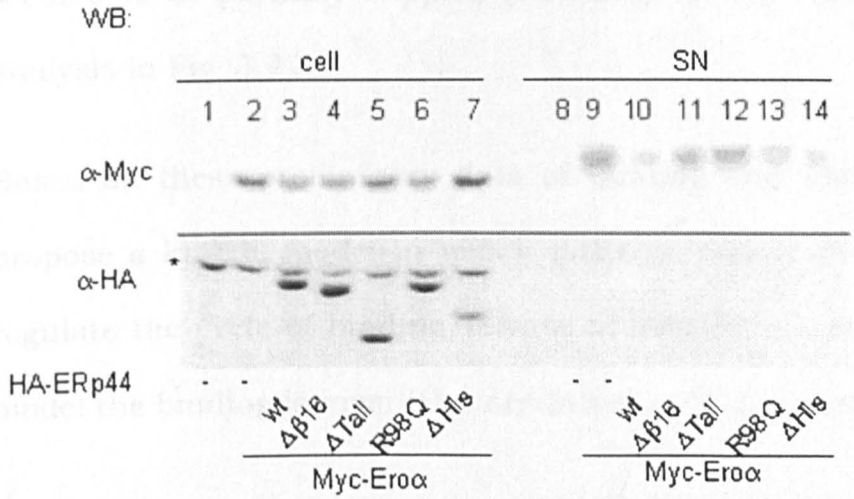


Figure 3.25. Role of C-tail on ERp44-mediated Ero1α retention

Supernatants after 4 h of culture in Optimem (SN) and cell lysates (cell) from HeLa cells transfected with Ero1α–Myc alone or Ero1α-Myc + HA-ERp44 mutants as indicated, were analyzed by Western blotting with monoclonal antibodies against Myc (Ero1α) or HA (ERp44). When over-expressed, Ero1α is partially secreted (lane 9). The over-expression of ERp44 reduces secretion to different extents (compare lanes 10-14). See Fig. 3.22 for the densitometric quantification of 4 to 14 experiments performed like the one showed in SDS-page, plotting the ratio between intracellular (IN) and secreted (OUT) Ero1α.

Then, we assessed the TMR activity of ERp44Δβ16 and ERp44ΔHis mutants in well-established functional assays in which over-expressed ERp44 retains over-expressed Ero1α (see part 3.4; Anelli et al., 2003; Otsu et al., 2006).

ERp44-mediated Ero1α retention assay showed that, although hypo-reactive, ERp44ΔHis displayed a substantially higher activity in Ero1α retention than wt (SDS-page in Fig. 3.25 and densitometric analysis in Fig. 3.22). Consistently with our model in Fig. 4.1, also Ero1α retention activity of ERp44ΔHis is weakened by NH_4Cl treatment (Fig. 3.22).

TMR activity of two hyper-reactive mutants, ERp44Δβ16 and ERp44ΔTail seemed to be partially crippled (SDS-page in Fig 3.25 and densitometric analysis in Fig. 3.22).

Based on these preliminary data of binding and TMR activity we could propose a kinetic model in which different region of C-tail contribute to regulate the cycle of binding/release of ERp44: according to our working model the binding is inversely correlated to TMR activity.

ERp44 should bind its substrates in ERGIC and release in ER, thereby trafficking along ESC should be well-regulated to support ESC folding demand. We have previously showed that ERGIC-53 contributes to dictate the ERGIC-localization of endogenous ERp44 (Anelli et al., 2007). Although over-expressed ERp44 wt localized mainly in the ER (Anelli et al., 2002), conserved TMR activity (Anelli et al., 2003) (Otsu et al., 2006) (Anelli et al., 2007) (Wang et al., 2007) (Fraldi et al., 2008) (Mariappan et al., 2008) (Cortini and Sitia 2010).

To study whether binding/release cycle (or at least the hyper-hypo reactivity of ERp44 mutants) is linked to localization/trafficking of the protein, we proceeded to the characterization of hyper-reactive and hypo-reactive mutants. Similarly to over-expressed wt protein, hyper-reactive

ERp44Δβ16 accumulates mainly in the ER at the steady state in fixed HeLa cells (Fig. 3.26a and 3.26b).

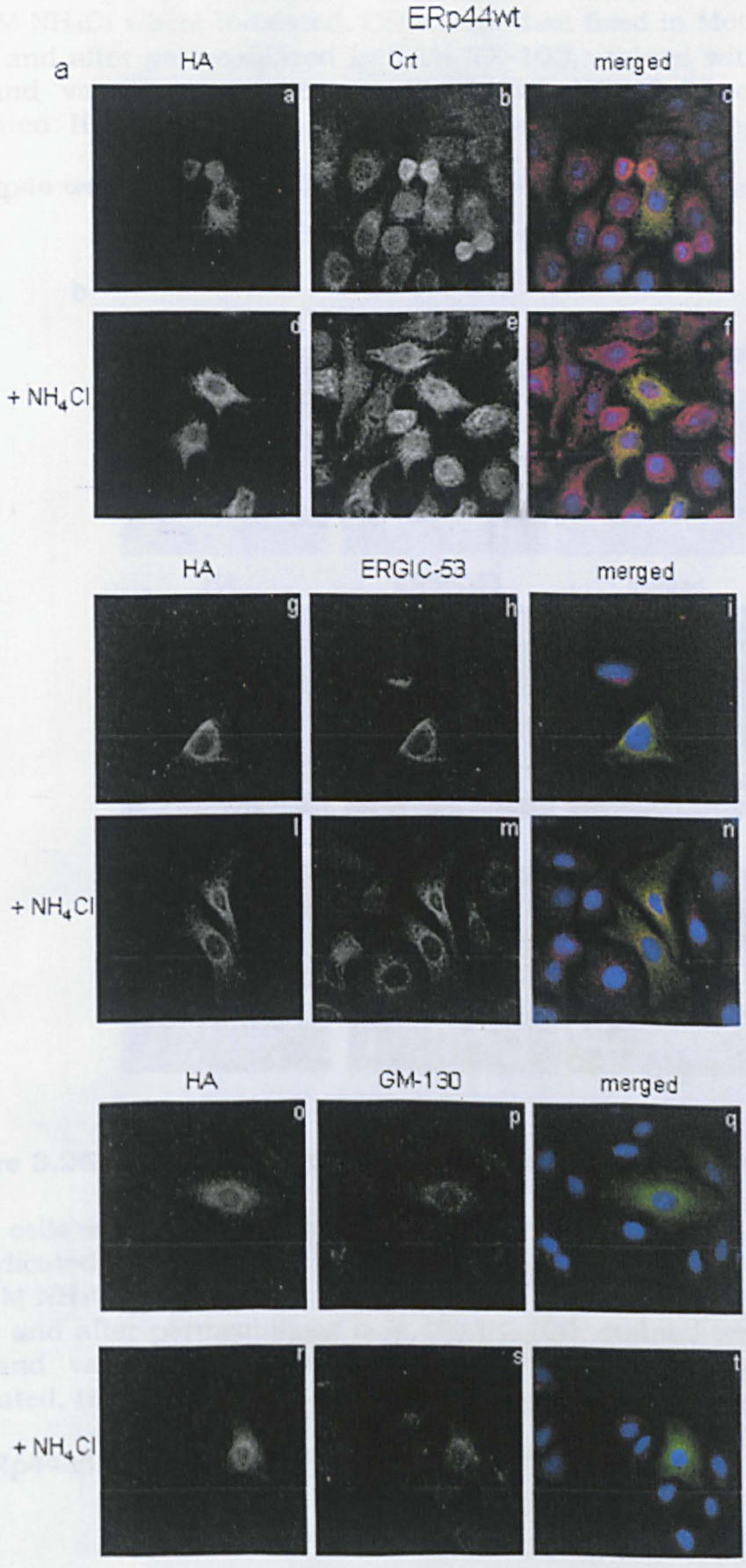


Figure 3.26a. ERp44ΔHis does not accumulate primarily in the ER

HeLa cells were cultured and transfected with HA-ERp44 wt, Δβ16 or ΔHis as indicated on cover slips. 48 h after transfection, cells were treated with 50 mM NH_4Cl where indicated. Cells were then fixed in Methanol 100% for 10 s. and after permeabilized in 0.1% TX-100, stained with antibodies α-HA and various organelles markers (Crt, ERGIC-53 or GM-130) as indicated. HA staining is in green, organelle marker staining in red.

A. ERp44 wt localized mainly in the ER also after NH_4Cl treatment.

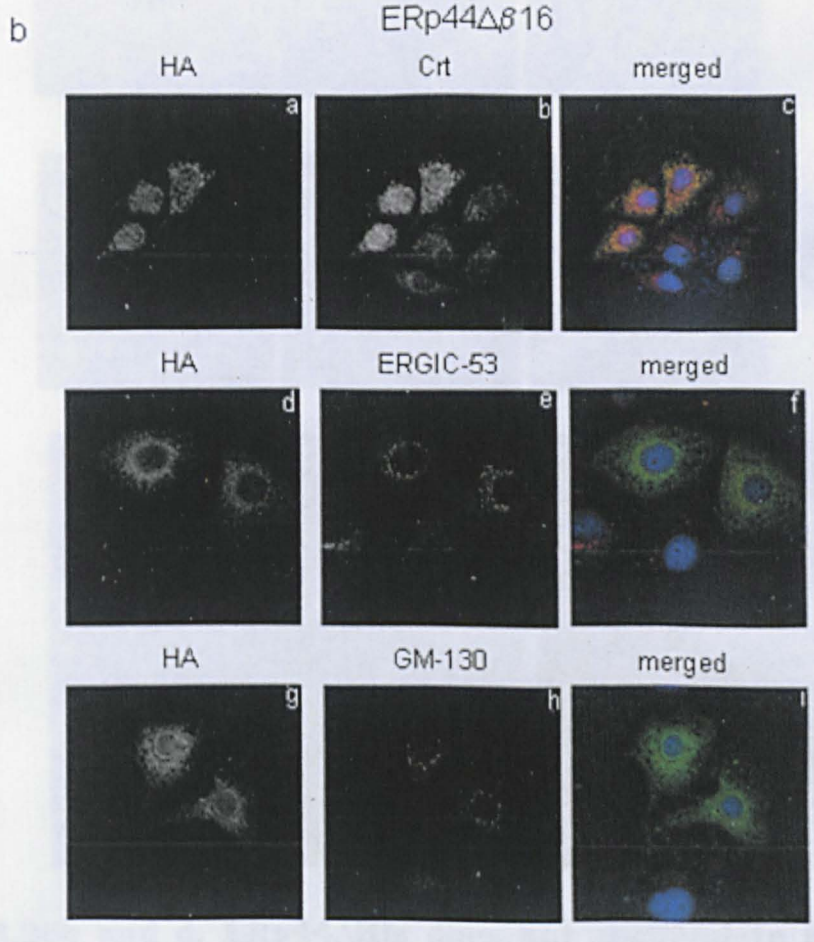


Figure 3.26b. ERp44ΔHis does not accumulate primarily in the ER

HeLa cells were cultured and transfected with HA-ERp44 wt, Δβ16 or ΔHis as indicated on cover slips. 48 h after transfection, cells were treated with 50 mM NH_4Cl where indicated. Cells were then fixed in Methanol 100% for 10 s. and after permeabilized in 0.1% TX-100, stained with antibodies α-HA and various organelles markers (Crt, ERGIC-53 or GM-130) as indicated. HA staining is in green, organelle marker staining in red.

B. ERp44Δβ16 localized mainly in the ER.

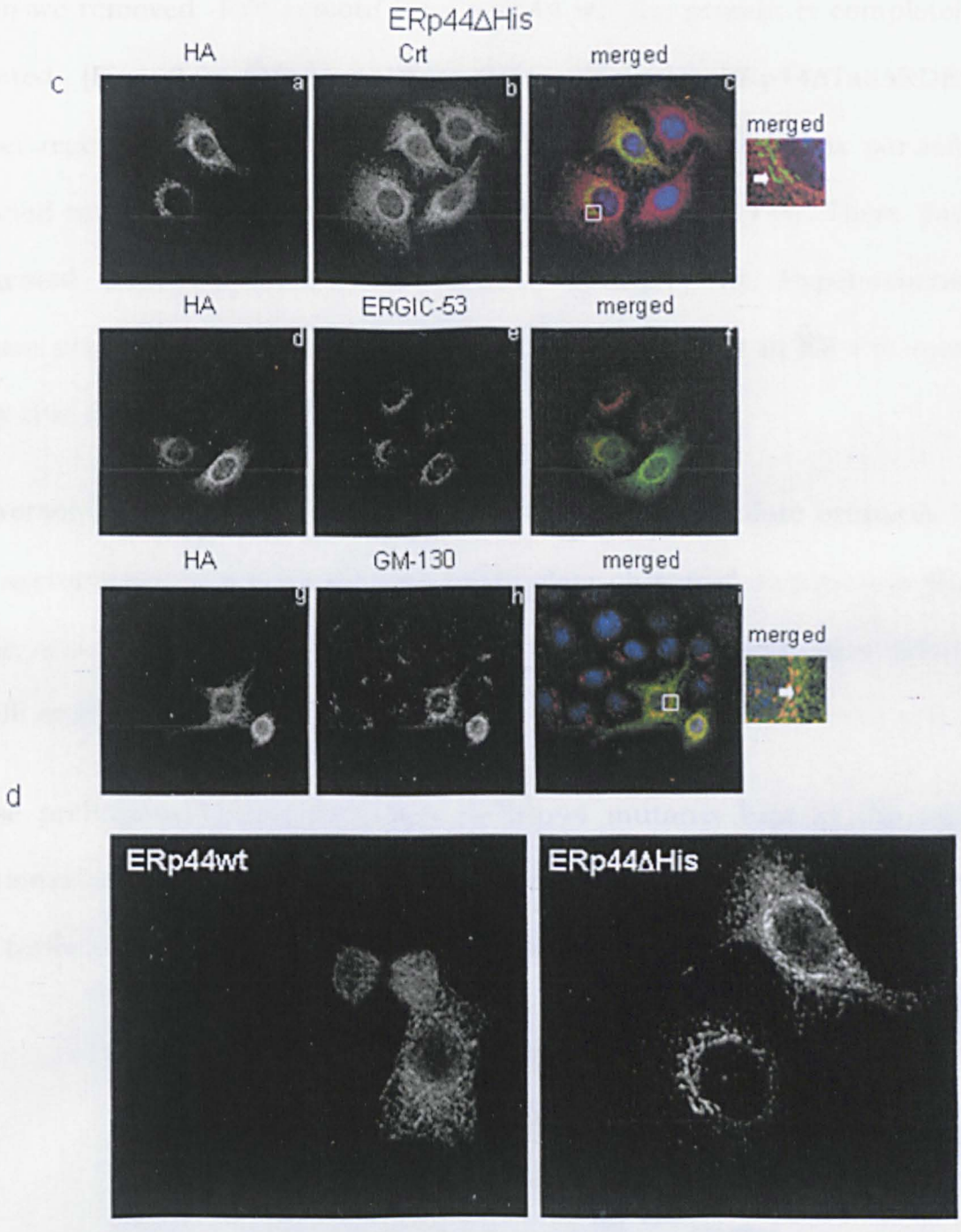


Figure 3.26c and d. ERp44ΔHis does not accumulate primarily in the ER

C. ERp44ΔHis localized does not accumulate mainly in the ER. For panels c and i, boxes show a higher magnification of the indicated area. Arrows indicate examples of co-localizing structures.

D. Magnification of panel a-Fig.A and panel a-Fig.C to compare the different distribution patterns of ERp44 wt to ΔHis.

When we removed -RDEL motif from ERp44 wt, the protein is completely secreted (Fig. 3.26-ERp44wtΔRDEL). By contrast ERp44ΔTailΔRDEL (hyper-reactive mutants in which -RDEL motif was removed) is partially retained intracellularly (Fig. 3.27- compare lanes 3-4 to 7-8). These data suggested that, although co-localized with ERp44 wt, hyper-reactive mutant are partially impaired in its trafficking or at least in ER exit most likely due to high affinity for ER matrix.

Conversely, ERp44ΔHis mutant does not seem to accumulate primarily in ER, assuming a partial asymmetric perinuclear clustered distribution (Fig. 3.25c and Fig 3.25d) overlapping the localization of endogenous ERp44 (Anelli et al., 2007) (Wang et al., 2007).

These preliminary characterization of ERp44 mutants hint at the tight relationship between kinetic cycle of binding/release (or al least reactivity) and trafficking that might modulate TMR activity of ERp44.

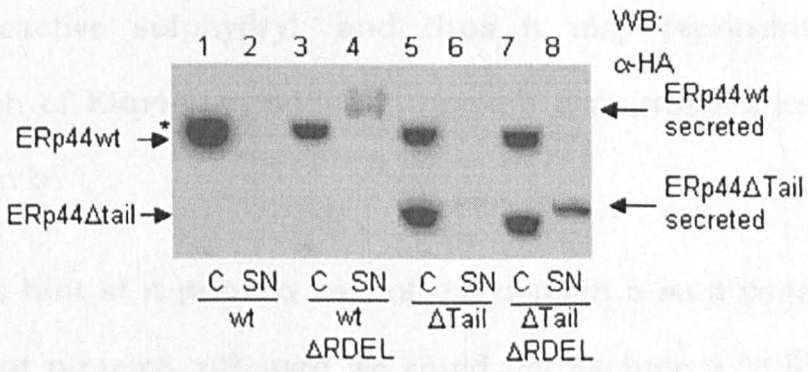


Figure 3.27. ERp44ΔTailΔRDEL is partially retained intracellularly

Supernatants after 16 h of culture in Optimem (SN) and cell lysates (C) from HeLa transfected with HA-ERp44 wt and mutants as indicated, were analyzed by western blotting with antibodies against HA (ERp44). ERp44 wtΔRDEL is fully secreted (lanes 3 and 4) while ΔTailΔRDEL is partially retained intracellularly (lanes 7 and 8). Asterisk marks anti-HA reactive background bands. Secreted ERp44wtΔRDEL and ΔTailΔRDEL are slower migrating than intracellular ones likely due to post-translational modifications (O-glycosylation) that take place beyond ER.

4. Discussion

4.1 ERp44 C-tail rearrangements

The function of many molecular chaperones depends on the regulation of the availability of Substrate Binding Site (SBS) by means of ATP-dependent conformational changes. Since ERp44 does not show any ATP binding nor hydrolysis, its function should be mediated, at least, through the transient nature of the contacts between C-tail and TRX-like domain a containing SBS. We showed that the destabilization of the interactions between C-tail and domain a (ERp44-T369A and ERp44Δβ16) and their strengthening (ERp44-T369C) modulated the binding activity of ERp44 suggesting a role of C-tail in regulating the accessibility of SBS during protein quality control in the Early Secretory Compartments (ESC)(Fig. 4.1). Indeed, MD simulations suggest that C-tail rearrangements (or at least the destruction of the interaction between β16 in the C-tail and the TRX domain a) trigger a conformational change in domain a suggestive of a highly reactive sulphydryl, and thus it may represent an active conformation of ERp44. In addition, domain a rearranges independently from domain b'.

The results hint at a primary role of the domain a as a primary binding site for client proteins although we could not exclude a "full" opening of the C-tail to expose also domain b' hydrophobic pocket providing an additional docking site to accommodate substrates. The inter-domain flexibility and the accessibility of SBS, both modulated by C-tail rearrangements seem to be the way whereby the ESC tunes ERp44 and thus, its protein quality control capacity.

The crystal structures of a prototype of PDI-family, yeast PDI (Tian et al., 2006) (Tian et al., 2008) revealed similar mechanisms of regulation of SBS accessibility. The spatial arrangement of the four domains in the homodimer (Tian et al., 2008) differed significantly from the structure reported for monomeric form (Tian et al., 2006). Moreover, in the dimer, both the active site in a' domain and the SBS in the b' domain are buried. This conformation of the dimer therefore presumably represents an "inactive" state of PDI, because most of its catalytic elements are inaccessible. On the other hand, monomeric state most likely represents an "active" form of PDI. The transition between monomeric and dimeric PDI could be a regulatory mechanism (Tian et al., 2008). Very recently, Freedman and co-workers reported another conformational change in human PDI (Van dat Nguyen et al., 2008) linked to SBS availability in domain b' very similar to conformational change describe in part 3.3 for ERp44. In both hPDI and yPDI, the b' domain is thought to be a primary SBS (Klappa et al., 1998)(Tian et al., 2006). Biophysical and NMR analysis revealed that the x region of PDI (see Fig. 1.6) can interact with the b' domain by "capping" a hydrophobic site on the b' domain. This site is most likely the SBS and hence such capping will inhibit substrate binding.

In addition to its role in modulating the accessibility of SBS of ERp44, the C-tail seems to be important for retrieval in the ESC. The "closed" conformation of the C-tail apparently could influence also the accessibility of -RDEL motif and therefore the interaction with its specific receptors (KDEL-R). Indeed, the retrieval of T369C mutant is partially rescued when a spacer (GFP) was employed to further protrude the -RDEL signal. An

intriguing possibility is that interaction with substrates exposes the -RDEL motif for binding to KDEL-R and retrieval.

Although retained intracellularly, ERp44-T369C-GFP does not form mixed disulfide with endogenous proteins suggesting that the Cys29-Cys369 disulfide bond locks the C-tail, hampering substrate binding. This observation is consistent with ERp44 T369C pull-down *in vitro* assays which shows that the full reactivity is restored only after DTT treatment. Hence, ERp44-T369C-GFP may cycle in ESC without any covalent interaction with substrates. Alternatively, the Cys29-Cys369 bond may be initially undone and then rapidly re-formed due to the local higher molar concentration of the C-Tail with respect to the endogenous substrates, *de facto* competing each other for the accessibility of SBS. According to this hypothesis a competition between the C-tail and substrates for SBS may be part of the ERp44 mechanisms of action.

Altogether, these data suggest that in living cells the C-tail rearrangements dynamically regulate both the reactivity of ERp44 with its endogenous client proteins, perhaps restricting its specificity, and the availability of the C-terminal –RDEL motif to KDEL-R (Fig 4.1).

It will be of interest to identify the mechanisms that control the ERp44 C-tail movements (see part 4.3)

4.2 ESC pH gradient-mediated regulation of ERp44

Part of the project was devoted to examine how ERp44-mediated Ero1α retention is modulated by ESC pH gradient. Based on the available

evidence, we could draw a unifying model of the ERGIC-53/ERp44 pathway, in which pH inversely controls the binding properties of ERGIC-53 and ERp44 (Fig. 4.1). The former is known to capture cargoes in the ER and release them in the Golgi in a pH-dependent way (Hauri et al., 2000) (Appenzeller-Herzog et al., 2004). According to our model, ERp44 captures its substrates in the *cis*Golgi, whose lower pH may modulate C-tail rearrangement to expose both the SBS and the -RDEL sequence. The ensuing complex, KDEL-R/ERp44/cargo, could be efficiently retrieved to the ER where it dissociates because of the higher pH.

The pH-mediated regulation of ERp44 may derive from disruption of the interface between domain a and the C-Tail, from interaction with other ligand partners, or a combination of both mechanisms. Indeed, the protonation of Arg98 in the ERGIC-*cis*Golgi complex might destabilize the interface between domain a and C-Tail through charge repulsion between the side chains of Arg98 and Arg367, thus favouring the C-tail rearrangements. In accordance with this hypothesis, the replacement of Arg with Gln partially inhibited both binding and TMR activity of the protein, likely because of the prevention of these interface-destabilizing repulsive interactions. ESC pH gradient may also regulate the protonation/deprotonation of residues involved in other functional states of ERp44 such as substrate binding/release and homodimer formation. A non-mutually exclusive hypothesis could be that the protonation may also destabilize the interface in the covalent homodimer. Indeed when we replaced Arg98 (located very close to Cys29 in domain a) with Gln, ERp44 accumulates as homodimer. This species is also present in ERp44 wt although to lesser extent suggesting that the mutation may capture a

physiological state of the protein. We have previously proposed (Anelli et al., 2003) that covalent homodimer of ERp44 may represent an intermediate species between monomeric and cargo-bound state. Since ERp44 does not display any of activity oxidase nor reductase (Wang et al., 2008), dimer form may itself provide oxidative power through disulfide exchange with substrates proteins and *vice versa.*

Although little is known about residues involved in homodimerization, Arg98 seems to play a key role in favouring the formation of this species. In accordance with our hypothesis NH_4Cl treatment induce a slight homodimerization in ERp44 wt partially mimicking R98Q mutation.

According to these data, the protonation of Arg98 may play a dual role in C-tail rearrangement and also in "homodimer regulation", destabilizing the interface between domain a and C-tail and between two ERp44 molecules. These two molecular mechanisms could be indeed two steps of the same mechanism of ERp44 activation in which Arg98 is the master regulator.

Other ATP-independent chaperones and folding helpers use ESC pH gradient to regulate their substrate binding and release. In RAP (receptor associated protein), an overall increase of the positive surface charge due to protonation of solvent-exposed histidine side chains in the Golgi complex support the dissociation of RAP chaperone from LPR (low-density lipoprotein receptor-related protein, member of the LDL receptor family- Lee et al., 2006). Histidine residues are part of interface between chaperone and substrate, and thus the histidine to alanine residue mutation impairs dissociation of RAP from LRP at lower pH. Similar studies (Appenzeller-Herzog et al., 2004) reported that His178 residue is

the hallmark of molecular pH sensor in ERGIC-53. In this case, upon arrival in the ERGIC, His-178 in ERGIC-53 is protonated because of lowered luminal pH, leading to the loss of one (or more) Calcium ions. Thereby, lectin activity of ERGIC-53 is inactivated and thus triggers cargo release. In living cell, neutralization of pH gradient in the ESC leads to inhibition of the release of ERGIC-specific cargo (Appenzeller-Herzog et al., 2004).

In these two cases histidine residues play a key role in binding/release cycle of the protein and thus, ESC pH gradient affect primarily the interaction with substrates.

Ero1α is the favourite substrate of ERp44 (Anelli et al., 2002) and when over-expressed is abundantly secreted. The co-over-expression of ERp44 fully retains intracellularly Ero1α. Moreover, Ero1α is not reported to be ERGIC-53 substrate (our unpublished data), a lectin inversely regulated by pH gradient (Appenzeller-Herzog et al., 2004). Being an ER resident protein, O-glycosylation and Golgi-glycan processing (both partially inhibited by NH_4Cl treatment) should not be important for its oxidative activity in ESC. IgM polymerization/secretion could be severely affected by NH_4Cl since they involve more complex quality control and processing machineries (Anelli et al., 2007). Hence, Ero1α is *bona fide* a good candidate as ERp44 TMR-substrate for our studies.

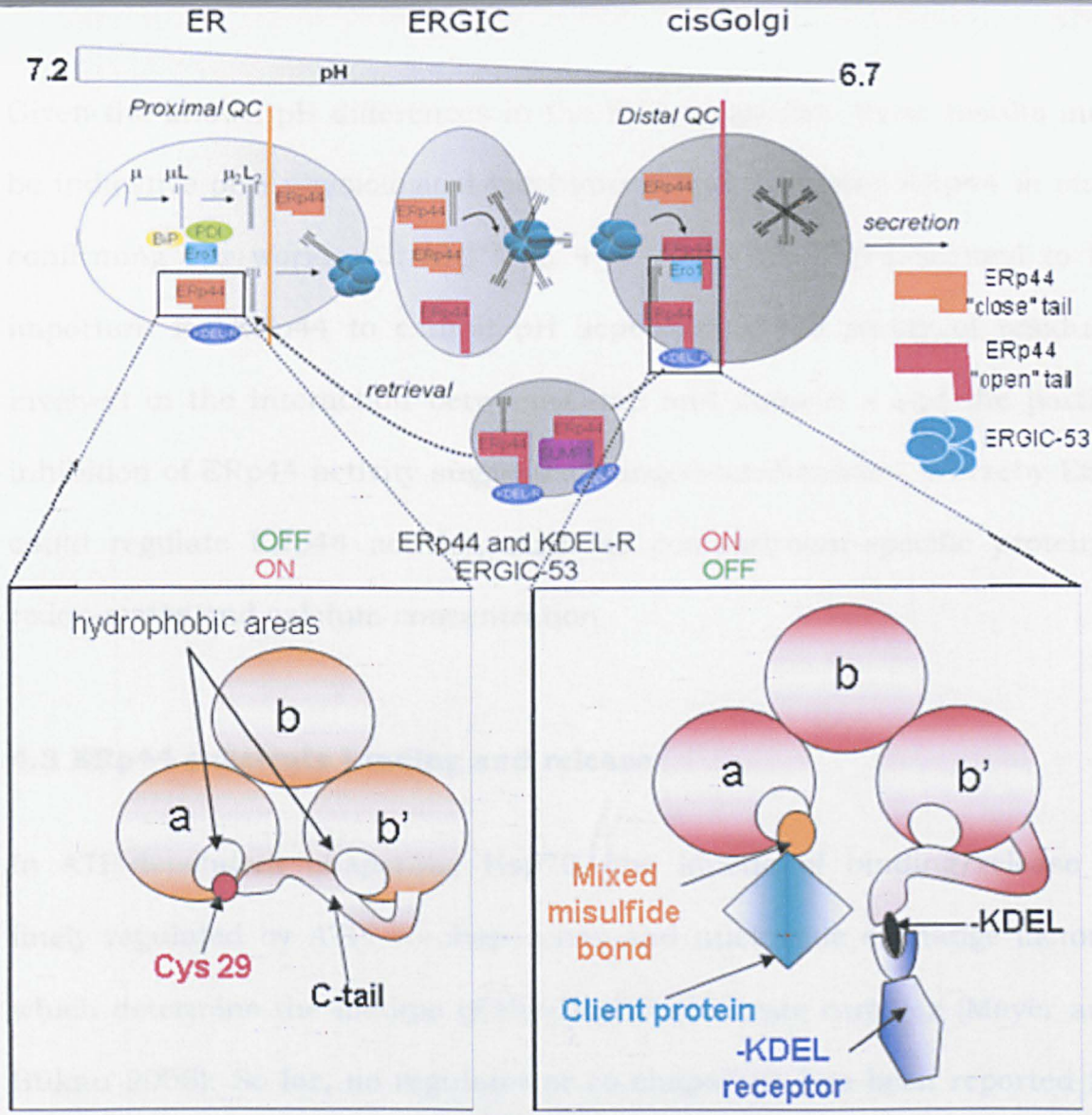


Figure 4.1. A possible model of the regulatory role of the C-tail in modulating SBS accessibility along ESC

The pH gradient between the ER and the Golgi complex (depicted as a shade of grey) could regulate ERp44 activity, by opening the C-tail (from orange-ERp44 to red-ERp44) simultaneously making the ERp44 SBS accessible to substrates (IgM subunits in grey, SUMF1 in purple or Ero1 in light blued) and the C-terminal RDEL motif to KDEL-Receptors (in blue). In the magnification panels, the cartoons depict clover-like structure of ERp44 where hydrophobic pockets (white) and the Cys29 (red ball) are shielded by the C-tail, yielding the close conformation (orange ERp44). Once the C-tail swings off to expose the SBS (Cys29 and hydrophobic patch surrounding in domain a) (red ERp44), client proteins could then interact with ERp44. The C-tail rearrangements also increase the exposure of -RDEL motif allowing the effective interaction with KDEL-R which in turn, retrieve via COPI vecicles the ERp44-cargo complex to the ER. Here, the C-tail swings onto the a domain, expelling the client protein, switching ERp44 to a less-active state (orange ERp44)

At the lower pH active ERp44 can retrieve unpolymerized IgM subunits, whilst native polymers are allowed to proceed further. Note the two-step IgM biogenesis in which the BiP/PDI-dependent μ_2L_2 assembly precedes the attempts to form polymers (dependent on ERGIC-53, ERp44 and other molecules (Anelli et al., 2007)(Anelli and Sitia 2008).

Given the known pH differences in the ESC organelles, these results may be indicative of a physiological mechanism that regulates ERp44 *in vivo*, confirming our working model (Fig. 4.1). Although Arg98 seemed to be important for ERp44 to exhibit pH dependence, the plenty of residues involved in the interaction between C-tail and domain a and the partial inhibition of ERp44 activity suggested hinged-mechanisms whereby ESC could regulate ERp44 activity, such as compartment-specific proteins, redox-states and calcium concentration.

4.3 ERp44 substrate binding and release

In ATP-dependent chaperone Hsp70, the kinetic of binding/release is finely regulated by ATP, co-chaperones and nucleotide exchange factors, which determine the lifetime of the Hsp70-substrate complex (Mayer and Bukau 2005). So far, no regulator or co-chaperone has been reported for ERp44.

Here, we show how different regions of C-tail contribute to regulate the kinetic of binding/release of ERp44 according to a kinetic model in which binding activity is inversely correlated to TMR activity.

Indeed, the removal of the last 9 residues of the C-tail (ERp44Δβ16) is sufficient to "unleash" the disulphide-mediated activity but cripples TMR activity. Moreover, the partial retention of ERp44ΔTailΔRDEL with respect to ERp44wtΔRDEL is highly suggestive of enhanced affinity for ER matrix, mirroring the hyper-reactivity of this class of mutants (ERp44ΔTail and ERp44Δβ16). According with these data, β16 strand could serve as a lid to drastically reduce the accessibility to client proteins of SBS, *de facto*

regulating binding/release of the protein. Thus, ERp44Δβ16 could bind to endogenous substrate in non-regulated way resulting in both ineffective interactions and impaired traffic along ESC. Although we reported that hyper-reactive ERp44ΔTail displayed an enhanced chaperone-like activity *in vitro* (see part 3.2), in living cells hyper-reactivity could reduce the "dynamic" properties of ERp44 (trafficking and substrate binding-release) resulting in less efficient TMR activity. ERp44Δβ16 behaviour may resemble the canonical ATP-dependent chaperones (Hsp70- Mayer and Bukau 2005) in which mutation in ATPase activity that cripples substrate release thus resulted in less efficient chaperone activity (Hendershot et al., 1995) (Wei et al.,1995).

On the other hand, the removal of His loop results in an enhanced efficiency of TMR activity of ERp44. Since His loop is remote to SBS in domain a, the removal of this region might affect ERp44 TMR likely independently from SBS-accessibility, as reported for ERp44Δβ16. Indeed, His loop may be involved in the modulation of ERp44, *per se* and/or by means of interaction with other macromolecular entities. In the crystal structure of ERp44, His-loop is invisible, likely due to its intrinsic disorder and/or flexibility. These peculiar features make His-loop a candidate as docking site for ligand proteins. Third party proteins might interact with this loop, regulating ERp44 activity according to ESC folding demand.

Third party proteins may inhibit ERp44 activity by means of slowing down the substrate-release, trafficking along ESC or a combination of both mechanisms. In this case, hypo-reactivity (or faster kinetic cycle of binding/release) is a consequence of missing interaction.

Alternatively, His-loop removal could enhance interaction with escort protein that meditated physiological ERGIC-localization of non-bound ERp44. However, both mechanisms resulted in a partial accumulation of ERp44 in downstream compartments.

Altogether, the enhanced TMR of ERp44ΔHis might reflect a enhanced "dynamic" properties of the mutant in terms of higher rate of turnover in the ESC and/or faster binding/release cycle. In this scenario, His loop may be considered as a "gear lever" by which ESC modulates ERp44 activity to support ESC folding demand. According with this hypothesis, the removal of His-loop might make ERp44 less tuned.

4.4 Concluding remarks

Altogether my findings indicate that the flexible C-terminal tail dynamically regulates both substrate binding and TMR activity of ERp44 during protein quality control in the ESC. Moreover, this work supports the hypothesis of a role of ESC pH gradient in governing ERp44-TMR activity, in which Arg98 seems to be a key residue in mediating pH sensitivity of ERp44.

A deeper knowledge of the structure/function relationship of ERp44 will shed light on the protein quality control mechanisms and thus provide essential knowledge of ESC processing diseases and in biotechnology, improving the production of man-made therapeutic proteins.

5. Material and methods

5.1 Cells and reagents

BL21 (DE3) and Rosetta-Gami 2 (DE3) were obtained from Novagen. Ni-NTA resin was purchased from Qiagen. Proteins were purified with AKTA purifier and column Superdex 200 10/30 (Amersham Pharmacia Biotech). Crystallization kits and plates were purchased from Hampton research.

HeLa, HEK293 and HepG2 cell lines were obtained from ATCC (Manassas, VA, USA), cultured in Dulbecco's modified Eagle's medium supplemented with 5% fetal calf serum (FCS), antibiotics and glutamine. FCS, Optimem and culture media were purchased from Gibco BRL; Sepharose-conjugated protein A and G, ECL reagents were purchased from Amersham-Pharmacia. DSP was purchased from Pierce.

Rabbit anti-PDI were kind gifts from Drs. R. M. E. Parkhouse (Pirbright Laboratory, Surrey, UK) and I. Braakman (Utrecht, NL); anti-ERp58 were generously provided by Dr. S. Bonatti (Naples,IT) , J. Saraste (Bergen, NO), E. Neve and R. Petterson (Stockholm, Sweden)(Neve et al., 2005). Rabbit anti-calreticulin was from Stressgen; mouse monoclonal anti-actin (clone AC-40) from Sigma Chemical Co; goat anti-mouse horseradish peroxidase (HRP); HRP rabbit anti-mouse IgG (H+L) was from DakoCytomation. Hoechst, FITC Goat anti-rat IgG (H+L), TRITC Goat anti-rabbit and mouse IgG (H+L) from Molecular Probes. Monoclonal Anti-GFP antibodies and Anti-His HRP conjugated were purchased from Santa Cruz

Mouse monoclonal antibodies specific for Myc (9E10) and HA (12CA5) were immobilized by cross-linking to Protein G and Protein A beads

respectively (Reddy et al., 1996). Lipofectin, Lipofectamine and Plus reagents were purchase from Invitrogen. Unless otherwise indicated, chemicals were purchased from Sigma Chemical Co.

5.2 Plasmids and vectors

pET28 vector was obtained from Novagen. The cDNA encoding human ERp44 without the signal sequence from pGEX-4T-1-ERp44 (Anelli et al., 2002) was amplified by PCR and cloned into pET28 vector at NheI and XhoI sites. The vectors for the expression of HA-ERp44 in mammalian cells were previously described (Anelli et al., 2002). HA-ERp44 mutants were obtained by PCR (PCR cloning in table 5.1) or by Site-Directed Mutagenesis (SDM in table 5.1). The PCR product is re-inserted in pcDNA3.1 (-) using the restriction sites XhoI–Acc65I. Construct was re-sequenced before transfection experiments. In order to retain the protein in the endoplasmic reticulum, HA-ERp44 mutants have the C-terminal RDEL motif left in place. The presence of the tag does not influence the function of ERp44 in any of the available assays (Anelli, Ceppi, Bergamelli and Sitia, unpublished results). Mouse version of ERp44-GFP-KDEL was a kind gift of Prof. E. Snapp (Albert Einstein College of Medicine, NY, USA).

	PRIMERS	PLASMI	TEMPLATE		ERp44
FW	CACCGCTAGCGAAATAACAAGTCTTGATACAGAGAAT	pET28	pGEX ERp44	PCR	ERp44w t
Rv	GCGCCTCGAGTTAATCCCTCAATAGAGTATACCTATAT			clonig	
FW	CACCGCTAGCGAAATAACAAGTCTTGATACAGAGAAT	pET28	pET28 ERp44	SDM	ERp44Δtail
Rv	GCGCCTCGAGTTAAGAATGTAAGTCAAATACGAATTG				
FW	GCACCCAGTGAATATAGGTATTGTCTATTGAGGGATTAACTCG	pET28	pET28 ERp44	SDM	ERp44 t369c
Rv	CGAGTTAATCCCTCAATAGACAATACCTATATTCACTGGGTGC				
FW	CACCGCTAGCGAAATAACAAGTCTTGATACAGAGAAT	pET28	pCDNA ERp44	PCR	ERp44
Rv	GCGCCTCGAGTTAATCCCTCAATAGAGTATACCTATAT		c29s c63s	cloning	c29s c63s
FW	CACCGCTAGCGACCCCATTCAAGAAATTC	pET28	pET28 ERp44	PCR	ERp44ΔTRX
Rv	GCGCCTCGAGTTAATCCCTCAATAGAGTATACCTATAT			cloning	
FW	GAGAATACAGGGGTCAGCAGTCAGTGAAAGCATTGGC	pCDNA	pCDNA ERp44	SDM	ERp44 r98q
Rv	GCCAATGCTTTCACTGACTGCTGACCCCTGTATTCTC				
FW	CCAGTGAATATAGGTATGCTCTATTGAGGGATCG	pCDNA	pCDNA ERp44	SDM	ERp44 t369a
Rv	CGATCCCTCAATAGAGCATACCTATATTCACTGG				
FW	CAACTCGAGCGTTACCATGCATCCTGCC	pCDNA	pCDNA ERp44	PCR	ERp44Δtail
Rv	CACGGTACCTTAAAGCTCATCTCGTTCTCTGTGCAGTTTTCCAGAA			cloning	
FW	GAAAACTGCACAGAGAATAACATCATGGACCTGACCC	pCDNA	pCDNA ERp44	SDM	ERp44Δtail
Rv	GGGTCAGGTCCATGATGTTATTCTCTGTGCAGTTTTC				ΔRDEL
FW	CAACTCGAGCGTTACCATGCATCCTGCC	pCDNA	pCDNA ERp44	PCR	ERp44Δβ16
Rv	GCGCGGTACCTTAAAGCTCATCTCGACTGGGTGCTAGTTTCTGGAAG			cloning	
FW1	GCGCGCGGATCCGGTAGTCCACCTGAGAGCTCCTTC	pCDNA	pCDNA ERp44	cut	ERp44Δhis
Rv1	GCGCGGATCCAGATCCAGAATGTAAGTCAAATACGAA			and	
FW	CAACTCGAGCGTTACCATGCATCCTGCC			paste	
Rv	GCGCGCGGTACCTTAAAGCTCATCTCGATCCCTC				
FW	CACCCAGTGAATATAGGTATTGTCTATTGAGGGATCGAGATG	pCDNA	pCDNA ERp44	SDM	ERp44 t369c
Rv	CATCTCGATCCCTCAATAGACAATACCTATATTCACTGGGTG				
FW	CCCAGCGAGTATAGGTATTGTTTATTGAGGGATCGAG	pEGFP N1	pEGFP N1	SDM	ERp44 t369c
Rv	CTCGATCCCTCAATAAACAATACCTATACTCGCTGGG		ERp44		

Table 5.1. ERp44 cloning primers.

5.3 Preliminary bioinformatic analysis

The bioinformatic analysis was carried on to identify physic-chemical characters, secondary structure and disordered regions. Several tools from the World Wide Web were used to analyze protein sequence of human ERp44.

The sequence was analysed through the following tools:

- *Prot. Param.*, a tool that allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL databases or for a user entered sequence. The computed parameters include the molecular weight, theoretical isoeletric point, amino acid composition, atomic composition,

extinction coefficient, estimated half-life, instability index. www.expasy.org.

- *RONN (Regional Order Neural Network),* a computational tool for prediction of disordered/unstructured regions within a protein sequence. http://www.strubi.ox.ac.uk/RONN (Yang et al. 2005).

- *CDD (Conserved Domain Database),* protein annotation resource that consists of a collection of well-annotated multiple sequence alignment models for ancient domains and full-length proteins. CDD content includes NCBI-curated domains, which use 3D-structure information to explicitly define domain boundaries and provide insights into sequence/structure/function relationships, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM). http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml (Marchler-Bauer et al., 2005).

5.4 Protein expression and purification

The pET28-ERp44 and mutants were expressed in BL21 (DE3) or Rosetta-Gami 2(DE3) competent cells for protein production. Cells were grown at 37°C in LB medium containing 50 mg/ml kanamycin. Protein expression was induced at a culture OD_{600} of 0.6/0.8 by addition of 0.1 mM IPTG. Cells were harvested after 18 h at 20°C by centrifugation and re-

suspended in 1/10th volume of a ice cold buffer containing PBS buffer + 1 mg/ml lysozyme, 20 mg/ml ribonuclease A, 2 mg/ml deoxyribonuclease I, Complete EDTA-free protease inhibitor cocktail and 2β-mercaptoethanol where indicated. Cells were disrupted by sonication at 4°C using a Bandelin sonicator. Cellular debris was removed by centrifugation at 25 000g for 30 min at 4°C.

5.4.1 Affinity purification

For a good balance of protein recovery versus specificity, 500 µl of Ni-NTA resin were employed for every 1 L of starting culture volume. Ni-NTA resin was added to the clear supernatant and incubated for 1 h with gentle shaking at 4°C and then loaded onto a polypropylene column. Unbound proteins were removed by washing the resin with 100 column volumes of washing buffer (PBS buffer + 15 mM imidazole). Bound proteins were eluted with the same buffer containing 0.5 M imidazole. The elution was repeated four times to maximize the recovery without diluting excessively the protein concentration.

5.4.2 Size exclusion chromatography

The affinity-purified ERp44 was injected onto a Superdex 200 10/30 size exclusion chromatography column equilibrated with 20 mM HEPES pH 7.5, 150 mM NaCl, 1mM DTT where necessary. Purified ERp44 was digested overnight at room temperature with thrombin at a ratio 1:1000 (w:w) after dialysis against a buffer containing 20 mM Tris-Cl pH 8.3, 150 mM NaCl, 2.5 mM EDTA. The digested protein was loaded onto a Superdex 200 chromatography column as described above to remove the

protease and the cleaved His tag. All mutated proteins were expressed and purified using the same protocol for the wild type ERp44 (except where indicated).

5.4.3 Selenomethionyl ERp44

Single-wavelength anomalous dispersion (SAD) is a technique used in X-ray crystallography that facilitates the determination of the structure of proteins or other biological macromolecules by allowing the solution of the phase problem. The most popular method of incorporating anomalous scattering atoms into proteins is to express the protein in a media rich in selenomethionine, which contains selenium atoms. The selenomethionyl ERp44 was prepared using the method of methionine-biosynthesis pathway inhibition (Doublié 1997) (Liu et al., 2005) and purified using the same protocol for the native ERp44 adding 5mM DTT in every purification step. The molecular weight of the recombinant ERp44 and the SeMet incorporation were determined by MALDI–TOF mass spectrometry.

5.4.4 Dynamic light scattering

Dynamic Light Scattering (DLS) is used to determine the size distribution profile of small particles in solution. In this instrument the sample is analysed based on the diffraction pattern of a laser light.

Protein solutions were centrifuged at 20000g for 30 min at 4°C to remove precipitates or dust particles. DLS measurements were performed at 4°C, 25°C and 37°C on a DynaPro MS/X instrument with temperature control (Protein Solutions). Laser intensity, time of acquisition and number of measurements were optimized to minimize the residual to the fitted

scattering curves. The ERp44 concentration used was 0.1 mg/ml. Data collection and deconvolution was performed using the DYNAMICSv.6 software. The degree of polydispersity (the lower the better) obtained from DLS measures were related to the probability to obtain crystals.

5.5 Crystallization and data collection

The crystallization process consists of two major events, nucleation and crystal growth. Nucleation is the step where the solute molecules dispersed in the solvent start to gather into clusters that can dissolve or become stable nuclei. It is at the stage of nucleation that the atoms arrange in a defined and periodic manner that defines the crystal structure. The crystal growth is the subsequent growth of the nuclei that succeed in achieving the critical cluster size. Nucleation and growth continue to occur simultaneously while the supersaturation exists. Supersaturation is a thermodynamically unstable state in which the solution contains more molecules dissolved than it would contain under the equilibrium. The supersaturation state can be further divided into three different zones: the metastable, the nucleation and the precipitation zone. Only in the nucleation zone the protein crystals nucleate and grow, while in the first one the solution may not nucleate for a long time, even if preformed crystals growth would be sustained, and in the precipitation zone proteins do not nucleate but precipitate out of solution. To supersaturate a protein solution there are two main possibilities: increasing the protein concentration, or adding a second reagent that reduces the sample solubility (Fig 5.1).

ERp44 was crystallized using mainly the hanging drop vapour-diffusion technique. This is the most used technique for the crystallization of macromolecules in which a drop composed of a mixture of sample and reagent is placed in vapor equilibration with a liquid reservoir of reagent. The drop contains a lower concentration of reagent than the reservoir, so water vapor slowly leaves the drop to achieve equilibrium, eventually ending up in the reservoir. As water leaves the drop, the sample undergoes an increase in relative supersaturation. Both the sample and reagent increase in concentration as water leaves the drop for the reservoir. Equilibration is reached when the reagent concentration in the drop is approximately the same as that in the reservoir.

Equal volumes of a protein solution at 10 mg/ml in 20 mM HEPES buffer pH 7.5, 150 mM NaCl and 1 mM DTT and precipitant solution (2 μl each) were mixed and placed on a cover slip over 500 ml of reservoir solution and incubated at 20°C. The crystals were dehydrated for ~24 h, then were transferred to artificial mother liquor containing 25% glycerol for cryoprotection and finally were immediately flash-cooled in a nitrogen stream at 100 K. Data collection at the wavelength of 0.97372Å at 100K was performed at the beam line ID29 of the European Synchrotron Radiation Facility Grenoble (France).

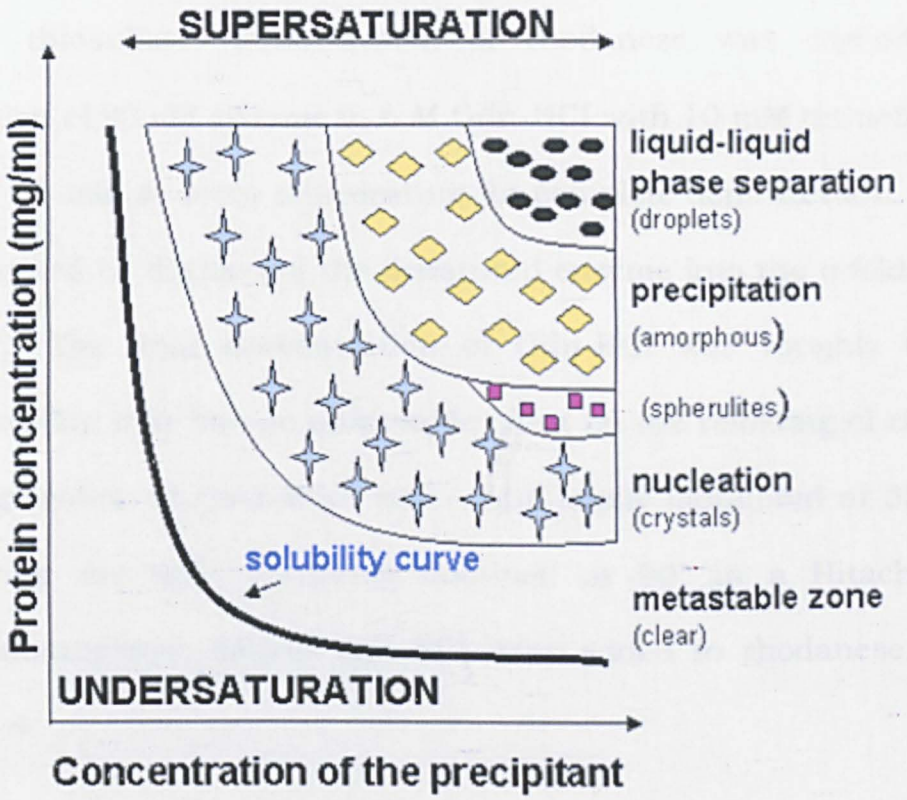


Figure 5.1 Solubility curve of a protein

5.6 Biochemistry and cell biology analysis

5.6.1 Pull-down assay

Pull-down assay of ERp44 proteins. Glutathiol Sepharose 4 fast-flow beads (GE Healthcare) were incubated at 25°C and pH 7.5 for 1 h with 10^{-6} M GST-Ero1α and various ERp44 proteins, as indicated, and then analysed by SDS-page under reducing conditions after being washed five times.

5.6.2 Chaperone-like activity

Chaperone activity was determined as described previously (Song and Wang 1995).Briefly, in all experiments, 0.2 M sodium phosphate, pH 7.5, was employed and the refolding buffer contained 5 mM DTT and 50 mM sodium thiosulfate. Denaturation of rhodanese was carried out by incubation of 90 μM enzyme in 6 M Gdn-HCl with 10 mM dithiothreitol for at least 30 min at room temperature for complete denaturation. Refolding was initiated by dilution of the denatured enzyme into the refolding buffer at 20°C. The final concentration of Gdn-HCl was roughly 0.3 M, a concentration that has no observable effect on the refolding of rhodanese. The aggregation of rhodanese was continuously monitored at 350 nm by measuring the light scattering obtained at 90° in a Hitachi F-4010 spectrofluorometer. ERp44 and PDI were added to rhodanese at molar ratio of 8.

5.6.3 Western blots

HeLa, HEK293 and HepG2 cell lines were transfected and analysed as described previously (Anelli et al., 2002, 2003). Anti-HA antibodies (12CA5) used 1:1000 dilutions. Anti-Myc antibodies (9E10) 1:1000 dilution. Anti-GFP (1:500 dilution). Anti-His HRP conjugated (1:1000).

5.6.4 Chemical cross linking

Cells were washed and incubated for 30 min at 4°C with 1mM DSP on a rocking platform. The reaction was quenched by rinsing the cells once with 20mM Tris–HCl pH 7.4, followed by two incubations for 15 min at

room temperature with the same buffer. Cells were then lysed in RIPA buffer+ 10mM NEM+ complete™.

5.6.5 Immunoprecipitation

For isolation of ERp44/Ero1α complexes and ERp44/μΔCH1), aliquots of the cell lysates (corresponding to 10^6 or $3*10^6$ cells) were IPed with anti-HA or anti-Myc, immobilized on Protein A beads (Reddy & Corley 1998) washed three times with STN buffer (20 mM Tris-Cl pH 7.7, 150 mM NaCl , 0.25% NP-40).

For isolation of ERp44/ERGIC-53 complexes, cells were washed with PBS + 2 mM $CaCl_2$ and then cross-linked with DSP.

Cell lysates were IP with anti HA-antibodies immobilized on Protein A beads, and then washed three times with STN buffer (20 mM Tris pH 7.7, 150 mM NaCl, 0.25% NP-40)

5.6.6 Stability assay

Two days after transfection, HeLa cells were treated with cycloheximide 0.5 mM (CHX) for 1, 2 or 4 h as indicated. After treatment, cells were harvested and lysated in RIPA buffer (20 mM Tris pH 7.7, 150 mM NaCl, 0.1% SDS) + 10 mM NEM+ Complete™.

For detergent solubility assay, HeLa cells were lysed in NP-40 buffer (20 mM Tris-Cl pH 7.4, 150 mM NaCl, 1% NP-40) or SDS buffer (20 mM Tris pH 7.4, 150 mM NaCl, 2% SDS) at 48 h after transfection. NP-40 soluble and insoluble fractions were separated by centrifugation, and aliquots corresponding to the same number of cells resolved by SDS-page.

5.6.7 Secretion assay

For secretion assays, two days after transfection, cells were washed 3 times and then cultured in OPTIMEM, in presence of 50 mM NH_4Cl where indicated. After 4 to 16 hours, 10 mM NEM was added to block disulphide interchange, supernatants harvested and cells lysed in RIPA buffer (20 mM Tris-Cl pH 7.7, 150 mM NaCl, 0.1% SDS) + 10 mM NEM+ Complete™. (Anelli et al., 2002). Aliquots of supernatants (corresponding to $2*10^6$ cells) were precipitated with Sepharose-bound Concanavalin A to concentrate glycoproteins or IP with Sepharose-bound anti-HA antibodies or precipitated with 20% of TCA. Lysates of 10^5 cells and supernatants were resolved under reducing or non-reducing conditions as indicated

Western Blots images were acquired with the Chemidoc-it Imaging System (UVP). Densitometric analysis was performed using Image Quant 5.2. ERp44-meditated Ero1α retention was determined as the ratio between the signal revealed with anti-Myc antibodies in the intracellular fractions and the signal of the supernatant.

5.6.8 Immunofluorescence

HeLa cells were fixed with Methanol 100% for 10 seconds, and then stained as described in (Anelli et al., 2007) with HA-antibodies (rat 1:1000 dilution), Calreticulin (rabbit 1:500 dilution), ERGIC-53 or GM130 (mouse 1:500 dilution).

HeLa cells were then analyzed on an Olympus inverted fluorescence microscope (model IX70) with DeltaVision RT Deconvolution System

(Alembic, HSR, Milano). After deconvolution, images were processed with Adobe Photoshop 7.0 (Adobe Systems Inc.).

5.7 Molecular dynamics simulation

The X-ray structure of ERp44 was used to prepare the starting configuration for the Molecular Dynamics (MD) simulations of the wild type protein. The Δβ16-ERp44 structure was generated by deleting the coordinates of the Glu365-Arg372 segment. The missing loop in the a domain (Phe50-Glu53) was modelled with the ModLoop server (Fiser et al. 2000). The resulting structure was immersed in a cubic box (~ 97 x 97 x 97 $Å^3$) of ~ 29,000 TIP3P water molecules and 13 Na^+ counterions. All the simulations and the subsequent analysis were performed with GROMACS 3.3.3., using the ff-amber99sb porting (Sorin & Pande 2005) of the AMBER parm99SB parameter set (Hornak et al. 2006).

Periodic boundary conditions were imposed. The equations of motion were integrated using the leap-frog method with a 1-fs timestep. The Berendsen algorithm was employed for temperature (T=300 K) and pressure (p=1 bar) regulation, with coupling constants of 0.1 ps. Bonds to hydrogen atoms were frozen with the LINCS method for the protein and the ligand, while SETTLE (was used for water molecules. The Particle Mesh Method was used to calculate electrostatic interactions, with a 11-Å cutoff for the direct space sums, a 1.0-Å FFT grid spacing and a 6-order interpolation polynomial for the reciprocal space sums. For van der Waals interactions, a switching function was used with a double 9-10 Å cutoff. Long-range

corrections to the dispersion energy were also included (Shirts et al., 2007).

The system was first minimized with 2000 steps of steepest descent. Harmonic positional restraints (with a force constant of ~ 12 kcal/mol/ $Å^2$) were then imposed onto the protein heavy atoms and gradually turned off in 400 ps, while the temperature was increased from 200 to 300 K. The system was then simulated for 5 ns. To assess the reproducibility of the results, two trajectories were generated for the wild type protein following the same protocol. The Δβ16 protein simulation was prolonged to 10 ns.

6. REFERENCES

Alberini CM, Bet P, Milstein C, Sitia R.,1990. Secretion of immunoglobulin M assembly intermediates in the presence of reducing agents. Nature. Oct 4;347(6292):485-7

Anelli, T. & Sitia, R., 2008. Protein quality control in the early secretory pathway. The EMBO journal, 27(2), 315-27.

Anelli, T. et al., 2002. ERp44, a novel endoplasmic reticulum folding assistant of the thioredoxin family. The EMBO journal, 21(4), 835-44.

Anelli, T. et al., 2007. Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. The EMBO journal, 26(19), 4177-88.

Anelli, T. et al., 2003. Thiol-mediated protein retention in the endoplasmic reticulum: the role of ERp44. The EMBO journal, 22(19), 5015-22.

Appenzeller, C. et al., 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. Nature cell biology, 1(6), 330-4..

Appenzeller-Herzog, C. & Ellgaard, L., 2008. The human PDI family: versatility packed into a single fold. Biochimica et biophysica acta, 1783(4), 535-48..

Appenzeller-Herzog, C. & Hauri, H., 2006. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. Journal of cell science, 119(Pt 11), 2173-83.

Appenzeller-Herzog, C. et al., 2004. pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. The Journal of biological chemistry, 279(13), 12943-50.

Aridor, M., 2007. Visiting the ER: the endoplasmic reticulum as a target for therapeutics in traffic related diseases. Advanced drug delivery reviews, 59(8), 759-81.

Awad, W. et al., 2008. BiP mutants that are unable to interact with endoplasmic reticulum DnaJ proteins provide insights into interdomain interactions in BiP. Proceedings of the National Academy of Sciences of the United States of America, 105(4), 1164-9.

Bannykh, S., Nishimura, N. & Balch, W., 1998. Getting into the Golgi. Trends in Cell Biology, 8(1), 21-25.

Barak, N.N. et al., 2009. Crystal structure and functional analysis of the protein disulfide isomerase-related protein ERp29. Journal of molecular biology, 385(5), 1630-42.

Becker, B. & Melkonian, M., 1996. The secretory pathway of protists: spatial and functional organization and evolution. Microbiological reviews, 60(4), 697-721.

Bertoli, G. et al., 2004. Two conserved cysteine triads in human Ero1alpha cooperate for efficient disulfide bond formation in the endoplasmic reticulum. The Journal of biological chemistry, 279(29), 30047-52.

Bezprozvanny, I., 2005. The inositol 1,4,5-trisphosphate receptors. Cell calcium, 38(3-4), 261-72.

Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF, Gething MJ.1993. Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell. Nov 19;75(4):717-28

Bobbert, T. et al., 2005. Changes of adiponectin oligomer composition by moderate weight reduction. Diabetes, 54(9), 2712-9.

Bole DG, Hendershot LM, Kearney JF., 1986. Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. J Cell Biol. May;102(5):1558-66.

Bonfanti, L. et al., 1998. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. Cell, 95(7), 993-1003.

Bonifacino JS, Suzuki CK, Lippincott-Schwartz J, Weissman AM, Klausner RD., 1989. Pre-Golgi degradation of newly synthesized T-cell antigen receptor chains: intrinsic sensitivity and the role of subunit assembly. J Cell Biol. Jul;109(1):73-83.

Boyadjiev, S.a. et al., 2006. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. Nature genetics, 38(10), 1192-7.

Braakman, I. & Sitia, R., 2003. Quality control in the endoplasmic reticulum protein factory. Nature, 426(6968), 891-4.

Brewer, J.W. & Hendershot, L.M., 2005. Building an antibody factory: a job for the unfolded protein response. Nature immunology, 6(1), 23-9.

Bukau, B., Weissman, J. & Horwich, A., 2006. Molecular chaperones and protein quality control. Cell, 125(3), 443-51.

Cabral, C.M., Liu, Y. & Sifers, R.N., 2001. Dissecting glycoprotein quality control in the secretory pathway. Trends Biochem. Sci, 26, 619-624.

Calfon, M. et al., 2002. IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. Nature, 415(6867), 92-6.

Caramelo, J.J. et al., 2004. The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates. The Journal of biological chemistry, 279(44), 46280-5.

Caramelo, J.J. et al., 2003. UDP-Glc:glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches in molten globule-like folding intermediates. Proceedings of the National Academy of Sciences of the United States of America, 100(1), 86-91.

Chandran, M. et al., 2003. Adiponectin: more than just another fat cell hormone? Diabetes care, 26(8), 2442-50.

Christianson, J.C. et al., 2008. OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. Nature cell biology, 10(3), 272-82.

Christis, C., Braakman, I. & Lubsen, N.H., 2008. Protein folding includes oligomerization - examples from the endoplasmic reticulum and cytosol. The FEBS journal, 275(19), 4700-27.

Coe, H. & Michalak, M., 2010. ERp57, a multifunctional endoplasmic reticulum resident oxidoreductase. The international journal of biochemistry & cell biology, 7-10.

Cortini, M. & Sitia, R., 2010. ERp44 and ERGIC-53 Synergize in Coupling Efficiency and Fidelity of IgM Polymerization and Secretion. Traffic (Copenhagen, Denmark)

Cosma, M.P. et al., 2003. The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. Cell, 113(4), 445-56..

Cunnea, P.M., 2002. ERdj5, an Endoplasmic Reticulum (ER)-resident Protein Containing DnaJ and Thioredoxin Domains, Is Expressed in Secretory Cells or following ER Stress. Journal of Biological Chemistry, 278(2), 1059-1066.

D'Arcy A., 1994. Crystallizing proteins - a rational approach?.Acta Crystallogr D Biol Crystallogr. Jul 1;50(Pt 4):469-71.

Danon, A., 2002. Redox reactions of regulatory proteins: do kinetics promote specificity? Trends in biochemical sciences, 27(4), 197-203.

Dean, N. & Pelham, H.R., 1990. Recycling of proteins from the Golgi compartment to the ER in yeast. The Journal of cell biology, 111(2), 369-77.

de Virgilio, M. et al., 1999. Degradation of a short-lived glycoprotein from the lumen of the endoplasmic reticulum: the role of N-linked glycans and the unfolded protein response. Molecular biology of the cell, 10(12), 4059-73.

Dias-Gunasekara, S. et al., 2005. Tissue-specific expression and dimerization of the endoplasmic reticulum oxidoreductase Ero1beta. The Journal of biological chemistry, 280(38), 33066-75.

Dierks, T. et al., 2005. Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme. Cell, 121(4), 541-52.

Dierks, T. et al., 2003. Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. Cell, 113(4), 435-44.

Dobson, C.M., 2003. Protein folding and misfolding. Nature, 426(6968), 884-90.

Dong, G. et al., 2009. Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. Immunity, 30(1), 21-32.

Doublié S., 1997. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. 276:523-3

Ellgaard, L. & Frickel, E., 2003. Calnexin, Calreticulin, and ERp57. Cell Biochemistry and Biophysics, 39(4).

Ellgaard, L. & Helenius, a., 2003. Quality control in the endoplasmic reticulum. Cell, 4(March).

Ferré-D'Amaré AR, Burley SK.,1994. Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. Structure. 1994 May 15;2(5):357-9. Structure 1994 Jun 15;2(6):56

Fink, A.L., 2005. Natively unfolded proteins. Current opinion in structural biology, 15(1), 35-41.

Fiser, a., Do, R.K. & Sali, a., 2000. Modeling of loops in protein structures. Protein science : a publication of the Protein Society, 9(9), 1753-73.

Foresti, O. et al., 2003. A Phaseolin Domain Involved Directly in Trimer Assembly Is a Determinant for Binding by the Chaperone BiP. Society, 15(October), 2464-2475.

Forster, M.L. et al., 2006. Protein disulfide isomerase-like proteins play opposing roles during retrotranslocation. The Journal of cell biology, 173(6), 853-9.

Fraldi, A. et al., 2008. Multistep, sequential control of the trafficking and function of the multiple sulfatase deficiency gene product, SUMF1 by PDI, ERGIC-53 and ERp44. Human molecular genetics, 17(17), 2610-21.

Frand, A.R. & Kaiser, C.A., 1999. Ero1p oxidizes protein disulfide isomerase in a pathway for disulfide bond formation in the endoplasmic reticulum. Mol. Cell, 4, 469-477.

Frand, A.R. & Kaiser, C.A., 1998. The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. Mol. Cell, 1, 161-170.

Frenkel, Z. et al., 2003. Endoplasmic reticulum-associated degradation of mammalian glycoproteins involves sugar chain trimming to Man6-5GlcNAc2. The Journal of biological chemistry, 278(36), 34119-24.

Frickel, E. et al., 2004. ERp57 is a multifunctional thiol-disulfide oxidoreductase. The Journal of biological chemistry, 279(18), 18277-87.

Gadsby, D.C., Vergani, P. & Csanády, L., 2006. The ABC protein turned chloride channel whose failure causes cystic fibrosis. Nature, 440(7083), 477-83.

Geiger, S.R. & Cramer, P., 2008. crystallization communications. , (April), 413-418.

Gess, B. et al., 2003. The cellular oxygen tension regulates expression of the endoplasmic oxidoreductase ERO1-Lalpha. European Journal of Biochemistry, 270(10), 2228-2235.

Gething MJ, Sambrook J., 1989.Protein folding and intracellular transport: studies on influenza virus haemagglutinin. Biochem Soc Symp.55:155-66

Gilchrist, A. et al., 2006. Quantitative proteomics analysis of the secretory pathway. Cell, 127(6), 1265-81.

Glick, B.S. & Nakano, A., 2009. Membrane traffic within the Golgi apparatus. Annual review of cell and developmental biology, 25, 113-32.

Gooptu, B. & Lomas, D.a., 2009. Conformational pathology of the serpins: themes, variations, and therapeutic strategies. Annual review of biochemistry, 78, 147-76.

Gray, S.G., Crowe, J. & Lawless, M.W., 2009. Hemochromatosis: as a conformational disorder. The international journal of biochemistry & cell biology, 41(11), 2094-7.

Gross, E. et al., 2006. Generating disulfides enzymatically: reaction products and electron acceptors of the endoplasmic reticulum thiol oxidase Ero1p. Proceedings of the National Academy of Sciences of the United States of America, 103(2), 299-304.

Guerin, M. & Parodi, A.J., 2003. The UDP-glucose:glycoprotein glucosyltransferase is organized in at least two tightly bound domains from yeast to mammals. The Journal of biological chemistry, 278(23),

20540-6. Hamilton, S.L., 2005. Ryanodine receptors. Cell calcium, 38(3-4), 253-60.

Haas IG, Wabl M.,1983. Immunoglobulin heavy chain binding protein. Nature. Nov 24-30;306(5941):387-9.

Hammond, C. & Helenius, a., 1994. Quality control in the secretory pathway: retention of a misfolded viral membrane glycoprotein involves cycling between the ER, intermediate compartment, and Golgi apparatus. The Journal of cell biology, 126(1), 41-52.

Hampton, R.Y., ER-associated degradation in protein quality control and. Current Opinion in Cell Biology, 476-482.

Harding, H.P. et al., 2000. Stress-Induced Gene Expression in Mammalian Cells. , 6, 1099-1108.

Hatahet, F. et al., 2009. Protein disulfide isomerase: a critical evaluation of its function in disulfide bond formation. Antioxidants & redox signaling, 11(11), 2807-50.

Hauri, H. et al., 2000. COMMENTARY ERGIC-53 and traffic in the secretory pathway. Traffic, 596, 587-596.

Hebert, D.N. & Molinari, M., 2007. In and out of the ER: protein folding, quality control, degradation, and related human diseases. Physiological reviews, 87(4), 1377-408.

Hebert, D.N., Garman, S.C. & Molinari, M., 2005. The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. Trends in cell biology, 15(7), 364-70.

Helenius A., 1994. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. Mol Biol Cell. Mar;5(3):253-65. Review

Helenius, a. & Aebi, M., 2004. Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem. 2004;73:1019-49.

Hendershot LM, Kearney JF., 1988.A role for human heavy chain binding protein in the developmental regulation of immunoglobin transport. Mol Immunol. Jun;25(6):585-95.

Hendershot LM, Wei JY, Gaut JR, Lawson B, Freiden PJ, Murti KG., 1995. In vivo expression of mammalian BiP ATPase mutants causes disruption of the endoplasmic reticulum.Mol Biol Cell. Mar;6(3):283-96.

Heras, B. & Martin, J.L., 2005. Post-crystallization treatments for improving diffraction quality of protein crystals. Acta crystallographica. Section D, Biological crystallography, 61(Pt 9), 1173-80.

Herscovics, A., Romero, P.A. & Tremblay, L.O., 2002. Glyco-Forum section XVII International Symposium on. Glycobiology, 12(4), 13-15.

Higo, T. et al., 2005. Subtype-specific and ER lumenal environment-dependent regulation of inositol 1,4,5-trisphosphate receptor type 1 by ERp44. Cell, 120(1), 85-98..

Hirao, K. et al., 2006. EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. The Journal of biological chemistry, 281(14), 9650-8.

Hobbs HH, Russell DW, Brown MS, Goldstein JL., 1990. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. Annu Rev Genet. ;24:133-70.

Hollien, J. & Weissman, J.S., 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. Science (New York, N.Y.), 313(5783), 104-7.

Holst B, Bruun AW, Kielland-Brandt MC, Winther JR 1996. Competition between folding and glycosylation in the endoplasmic reticulum. EMBO J. Jul 15;15(14):3538-46

Hornak, V. et al., 2006. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Bioinformatics, 725(May), 712-725.

Hurtley SM, Helenius A., 1989. Protein oligomerization in the endoplasmic reticulum. Annu Rev Cell Biol.;5:277-307.

Isidoro, C. et al., 1996. Exposed thiols confer localization in the endoplasmic reticulum by retention rather than retrieval. J. Biol. Chem, 271, 26138±26142.

Janiszewski, M. et al., 2005. Regulation of NAD(P)H oxidase by associated protein disulfide isomerase in vascular smooth muscle cells. The Journal of biological chemistry, 280(49), 40813-9.

Jansens, A., van Duijn, E. & Braakman, I., 2002. Coordinated nonvectorial folding in a newly synthesized multidomain protein. Science (New York, N.Y.), 298(5602), 2401-3.

Jiang, J. et al., 2005. Structural basis of interdomain communication in the Hsc70 chaperone. Molecular cell, 20(4), 513-24.

Jiang, Y. et al., 1997. Mapping of a gene for the increased susceptibility of B1 cells to Mott cell formation in murine autoimmune disease. Journal of immunology (Baltimore, Md. : 1950), 158(2), 992-7..

Jolliffe, N.a., Craddock, C.P. & Frigerio, L., 2005. Pathways for protein transport to seed storage vacuoles. Biochemical Society transactions, 33(Pt 5), 1016-8.

Karplus, M. & Kuriyan, J., 2005. Molecular dynamics and protein function. Proceedings of the National Academy of Sciences of the United States of America, 102(19), 6679-85.

Kaufman, R.J., 2002. Orchestrating the unfolded protein response in health and disease. Perspective, 110(10), 1389-1398.

Kaufman, R.J., 1999. coordination of gene transcriptional and translational controls Stress signaling from the lumen of the endoplasmic reticulum : coordination of gene transcriptional and translational controls. Genes & Development, 1211-1233.

Kawasaki, N. et al., 2008. The sugar-binding ability of ERGIC-53 is enhanced by its interaction with MCFD2. Blood, 111(4), 1972-9.

Kim SK, Kim YK, Lee AS., 1990. Expression of the glucose-regulated proteins (GRP94 and GRP78) in differentiated and undifferentiated mouse embryonic cells and the use of the GRP78 promoter as an expression system in embryonic cells. Differentiation. Feb;42(3):153-9.

Klappa, P. et al., 1998. The bJ domain provides the principal peptide-binding site of protein disul de isomerase but all domains contribute to binding of misfolded proteins. , 17(4), 927-935.

Klumperman, J. et al., 1998. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. Journal of cell science, 111 (Pt 2, 3411-25.

Kopito, R.R. & Sitia, R., 2000. Aggresomes and Russell bodies. EMBO Reports, 1(3), 225-231.

Koivu J, Myllylä R, Helaakoski T, Pihlajaniemi T, Tasanen K, Kivirikko KI,. 1987. A single polypeptide acts both as the beta subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase.J. Biol Chem. May 15;262(14):6447-9.

Kreis TE, Lodish HF., 1986. Oligomerization is essential for transport of vesicular stomatitis viral glycoprotein to the cell surface. Cell. Sep 12;46(6):929-37.

Kroczynska, B. et al., 2005. BIP co-chaperone MTJ1/ERDJ1 interacts with inter-alpha-trypsin inhibitor heavy chain 4. Biochemical and biophysical research communications, 338(3), 1467-77.

Lahtinen, U., Svensson, K. & Pettersson, R.F., 1999. Mapping of structural determinants for the oligomerization of p58, a lectin-like protein of the intermediate compartment and cis-Golgi. Eur. J. Biochem., 260, 392-397.

Lalla, C.D. & Sitia, R., Molecular Immunology. Molecular Immunology, 726-734.

Lander, E.S. et al., 2001. Initial sequencing and analysis of the human genome. Nature, 409(6822), 860-921.

Lappi, a. et al., 2004. A Conserved Arginine Plays a Role in the Catalytic Cycle of the Protein Disulphide Isomerases. Journal of Molecular Biology, 335(1), 283-295.

Lara-castro, C. et al., 2006. Syndrome Trait Cluster. In Vivo, 55(January), 249-259.

Lederkremer, G.Z., 2009. Glycoprotein folding, quality control and ER-associated degradation. Current Opinion in Structural Biology, (Figure 1), 515-523.

Li, Y. & Camacho, P., 2004. Ca2+-dependent redox modulation of SERCA 2b by ERp57. The Journal of cell biology, 164(1), 35-46.

Liebert, M.A. et al., 2006. Comprehensive Invited Review. Control, 8.

Liu, Z. et al., 2005. Salvaging Pyrococcus furiosus protein targets at SECSG. Journal of Structural and Functional Genomics, 121-127.

Luan, P. et al., 2000. A new functional domain of guanine nucleotide dissociation inhibitor (alpha-GDI) involved in Rab recycling. Traffic (Copenhagen, Denmark), 1(3), 270-81.

Luo, S. et al., 2006. GRP78/BiP is required for cell proliferation and protecting the inner cell mass from apoptosis during early mouse embryonic

Mainieri, D. et al., 2004. Zeolin. A New Recombinant Storage Protein Constructed Using Maize g-Zein and Bean Phaseolin 1. Society, 136(November), 3447-3456.

Marchler-Bauer, A. et al., 2005. CDD: a Conserved Domain Database for protein classification. Nucleic acids research, 33(Database issue), D192-6.

Mariappan, M. et al., 2008. ERp44 mediates a thiol-independent retention of formylglycine-generating enzyme in the endoplasmic reticulum. The Journal of biological chemistry, 283(10), 6375-83.

Masciarelli, S. & Sitia, R., 2008. Building and operating an antibody factory: redox control during B to plasma cell terminal differentiation. Biochimica et biophysica acta, 1783(4), 578-88.

Mast, S. & Moremen, K., 2006. Family 47 α□Mannosidases in N□Glycan Processing. Methods in Enzymology, 415(06), 31-46.

Mathieu, M.E. et al., 1996. Folding, unfolding, and refolding of the vesicular stomatitis virus glycoprotein. Biochemistry, 35(13), 4084-93.

Mattioli, L. et al., 2006. ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies. Journal of cell science, 119(Pt 12), 2532-41.

Mayer, M., Reinstein, J. & Buchner, J., 2003. Modulation of the ATPase Cycle of BiP by Peptides and Proteins. Journal of Molecular Biology, 330(1), 137-144.

Meunier L, Usherwood YK, Chung KT, Hendershot LM.,2002. A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. Mol Biol Cell. Dec;13(12):4456-69.

Meusser, B. et al., 2005. the long road to destruction. Nat. Cell Biol., 7, 766-772.

Mikoshiba, K., 2007. IP3 receptor/Ca2+ channel: from discovery to new signaling concepts. Journal of neurochemistry, 102(5), 1426-46.

Molinari, M. & Helenius, a., 1999. Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. Nature, 402(6757), 90-3.

Molinari, M. et al., 2003. Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. Science (New York, N.Y.), 299(5611), 1397-400.

Molinari, M. et al., 2002. Sequential assistance of molecular chaperones and transient formation of covalent complexes during protein degradation from the ER. The Journal of cell biology, 158(2), 247-57.

Molinari, M., 2000. Chaperone Selection During Glycoprotein Translocation into the Endoplasmic Reticulum. Science, 288(5464), 331-333.

Molinari, M., 2007. N-glycan structure dictates extension of protein folding or onset of disposal. Nature chemical biology, 3(6), 313-20..

Montecucco, C. & Molinari, M., 2006. Death of a chaperone. October, 443(October).

Munro S, Pelham HR., 1987. A C-terminal signal prevents secretion of luminal ER proteins.Cell. Mar 13;48(5):899-907

Nakatsukasa, K. & Brodsky, J.L., 2008. The recognition and retrotranslocation of misfolded proteins from the endoplasmic reticulum. Traffic (Copenhagen, Denmark), 9(6), 861-70..

Neve, E.P., Lahtinen, U. & Pettersson, R.F., 2005. Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. Journal of molecular biology, 354(3), 556-68..

Nguyen, V.D. et al., 2008. Alternative conformations of the x region of human protein disulphide-isomerase modulate exposure of the substrate binding b' domain. Journal of molecular biology, 383(5), 1144-55.

Nufer, O. et al., 2002. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. Journal of cell science, 115(Pt 3), 619-28.

Nyfeler, B. et al., 2006. Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. Traffic (Copenhagen, Denmark), 7(11), 1473-81.

Nyfeler, B. et al., 2008. Identification of ERGIC-53 as an intracellular transport receptor of alpha1-antitrypsin. The Journal of cell biology, 180(4), 705-12.

Nørgaard, P. et al., 2001. Functional differences in yeast protein disulfide isomerases. The Journal of cell biology, 152(3), 553-62.

Oda, Y. et al., 2003. EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. Science (New York, N.Y.), 299(5611), 1394-7.

Ohtsubo, K. & Marth, J.D., 2006. Glycosylation in cellular mechanisms of health and disease. Cell, 126(5), 855-67.

Okada, T. et al., 2002. Distinct roles of activating transcription factor 6 (ATF6) and double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK) in transcription during the mammalian unfolded protein response. The Biochemical journal, 366(Pt 2), 585-94.

Okuda-Shimizu, Y. & Hendershot, L.M., 2007. Characterization of an ERAD pathway for nonglycosylated BiP substrates, which require Herp. Molecular cell, 28(4), 544-54.

Oliver, J.D., 1997. Interaction of the Thiol-Dependent Reductase ERp57 with Nascent Glycoproteins. Science, 275(5296), 86-88.

Otsu, M. et al., Dynamic retention of Ero1alpha and Ero1beta in the endoplasmic reticulum by interactions with PDI and ERp44. Antioxidants & redox signaling, 8(3-4), 274-82.

Otte, S. & Barlowe, C., 2004. Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. Nat. Cell Biol., 6, 1189-1194.

Pagani, M. et al., 2000. Endoplasmic reticulum oxidoreductin 1-lbeta (ERO1-L), a human gene induced in the course of the unfolded protein response. J. Biol. Chem, 275, 23685-23692.

Peaper, D.R. & Cresswell, P., 2008. Regulation of MHC class I assembly and peptide binding. Annual review of cell and developmental biology, 24, 343-68.

Pelham HR, Munro S., 1993. Sorting of membrane proteins in the secretory pathway. Cell. Nov 19;75(4):603-5.

Pelham, H.R. & Rothman, J.E., 2000. The debate about transport in the Golgi--two sides of the same coin? Cell, 102(6), 713-9.

Periasamy, M. & Kalyanasundaram, A., 2007. SERCA pump isoforms: their role in calcium transport and disease. Muscle & nerve, 35(4), 430-42.

Pirot, P. et al., 2007. Global profiling of genes modified by endoplasmic reticulum stress in pancreatic beta cells reveals the early degradation of insulin mRNAs. Diabetologia, 50(5), 1006-14.

Pollard, M.G., Travers, K.J. & Weissman, J.S., 1998. Ero1p: a novel and ubiquitous protein with an essential role in oxidative protein folding in the endoplasmic reticulum. Mol. Cell, 1, 171-182.

Presley, J. F., Ward, T. H., Pfeifer, A. C., Siggia, E. D., Phair, R. D. and LippincottSchwartz, J. 1997. ER-to-Golgi transport visualized in living cells. Nature 389, 81- 417, 187-193.

Prostko CR, Brostrom MA, Brostrom CO., 1993. Reversible phosphorylation of eukaryotic initiation factor 2 alpha in response to endoplasmic reticular signaling. Mol Cell Biochem. Nov;127-128:255-65.

Pulvirenti, T. et al., 2008. A traffic-activated Golgi-based signalling circuit coordinates the secretory pathway. Nature cell biology, 10(8), 912-22.

Qiang, L., Wang, H. & Farmer, S.R., 2007. Adiponectin secretion is regulated by SIRT1 and the endoplasmic reticulum oxidoreductase Ero1-L alpha. Molecular and cellular biology, 27(13), 4698-707. Reddy, P.S. & Corley, R.B., 1998. Assembly, sorting, and exit of oligomeric proteins from the endoplasmic reticulum. BioEssays : news and reviews in molecular, cellular and developmental biology, 20(7), 546-54.

Reinhardt, C. et al., 2008. Protein disulfide isomerase acts as an injury response signal that enhances fibrin generation via tissue factor activation. , 118(3).

Ritter C, Helenius A. 2000. Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. Nat Struct Biol. Apr;7(4):278-80.

Rowe, S.M., Miller, S. & Sorscher, E.J., 2005. Cystic fibrosis. The New England journal of medicine, 352(19), 1992-2001.

Rueda, M. et al., 2007. A consensus view of protein dynamics. Proceedings of the National Academy of Sciences of the United States of America, 104(3), 796-801.

Russell, S.J. et al., 2004. The primary substrate binding site in the b' domain of ERp57 is adapted for endoplasmic reticulum lectin association. The Journal of biological chemistry, 279(18), 18861-9.

Sallese, M., Giannotta, M. & Luini, A., 2009. Coordination of the secretory compartments via inter-organelle signalling. Seminars in cell & developmental biology, 20(7), 801-9.

Sancho J, Chatila T, Wong RC, Hall C, Blumberg R, Alarcon B, Geha RS, Terhorst C., 1989. T-cell antigen receptor (TCR)-alpha/beta heterodimer formation is a prerequisite for association of CD3-zeta 2 into functionally competent TCR.CD3 complexes. J Biol Chem. Dec 5;264(34):20760-9.

Scales, S.J., Pepperkok, R. & Kreis, T.E., 1997. Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. Cell, 90, 1137-1148.

Scheckman RW.,1994. Regulation of membrane traffic in the secretory pathway.Harvey Lect. 1994-1995;90:41-57

Scherer, P.E. et al., 1995. to C1q , Produced Exclusively in Adipocytes *. Biochemistry, 26746-26749.

Selmer, T. et al., 1996. The evolutionary conservation of a novel protein modification, the conversion of cysteine to serinesemialdehyde in arylsulfatase from Volvox carteri. European journal of biochemistry / FEBS, 238(2), 341-5. Senderek, J. et al., 2005. Mutations in SIL1 cause Marinesco-Sjögren syndrome, a cerebellar ataxia with cataract and myopathy. Nature genetics, 37(12), 1312-4..

Sevier, C.S. & Kaiser, C.a., 2008. Ero1 and redox homeostasis in the endoplasmic reticulum. Biochimica et biophysica acta, 1783(4), 549-56.

Shen, Y. & Hendershot, L.M., 2005. ERdj3 , a Stress-inducible Endoplasmic Reticulum DnaJ Homologue , Serves as a CoFactor for BiP ' s Interactions with Unfolded Substrates. Molecular Biology of the Cell, 16(January), 40 -50.

Sitia R, Neuberger MS, Milstein C.,1987.Regulation of membrane IgM expression in secretory B cells: translational and post-translational events. EMBO J. Dec 20;6(13):3969-77.

Sitia R, Neuberger M, Alberini C, Bet P, Fra A, Valetti C, Williams G, Milstein C.,1990. Developmental regulation of IgM secretion: the role of the carboxy-terminal cysteine. Cell. Mar 9;60(5):781-90.

Sitia, R. et al., 2001. Manipulation of oxidative protein folding and PDI redox state in mammalian cells. The EMBO journal, 20(22), 6288-96.

Sitia, R. et al., 2000. The CXXCXXC motif determines the folding, structure and stability of human Ero1-Lalpha. The EMBO journal, 19(17), 4493-502.

Song, J. & Wang, C., 1995. in the refolding of rhodanese. , 316, 312-316.

Sorin, E.J. & Pande, V.S., 2005. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. Biophysical journal, 88(4), 2472-93.

Spreafico, M. & Peyvandi, F., 2008. Combined FV and FVIII deficiency. Haemophilia : the official journal of the World Federation of Hemophilia, 14(6), 1201-8.

Szegezdi, E. et al., 2006. Mediators of endoplasmic reticulum stress-induced apoptosis. EMBO reports, 7(9), 880-5.

Taylor, S.C. et al., 2004. The ER protein folding sensor UDP-glucose glycoprotein-glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. Nature Structural & Molecular Biology, 11(2), 128-134.

Tian, G. et al., 2008. The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule.

Tian, G. et al., 2006. The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. Cell, 124(1), 61-73.

Tu, B.P. & Weissman, J.S., 2004. Oxidative protein folding in eukaryotes: mechanisms and consequences. The Journal of cell biology, 164(3), 341-6.

Ungar, D., 2009. Golgi linked protein glycosylation and associated diseases. Seminars in cell & developmental biology, 20(7), 762-9

Ushioda, R. et al., 2008. ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. Science (New York, N.Y.), 321(5888), 569-72.

VENETIANER P, STRAUB FB., 1993.ENZYMIC FORMATION OF THE DISULFIDE BRIDGES OF RIBONUCLEASE. Acta Physiol Acad Sci Hung. 1963;24:41-53.

Vollenweider, F., 1998. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. J. Cell Biol., 142, 377-389.

Walker, K.W., Lyles, M.M. & Gilbert, H.F., 1996. Catalysis of oxidative protein folding by mutants of protein disulfide isomerase with a single active-site cysteine. Biochemistry, 35(6), 1972-80.

Walter, T.S. et al., 2006. Lysine Methylation as a Routine Ways & Means Rescue Strategy for Protein Crystallization. Bioinformatics, (November), 1617-1622.

Wang, L. et al., 2009. Reconstitution of human Ero1-Lalpha/protein-disulfide isomerase oxidative folding pathway in vitro. Position-dependent differences in role between the a and a' domains of protein-disulfide isomerase. The Journal of biological chemistry, 284(1), 199-206.

Wang, Z.V. et al., 2007. Secretion of the adipocyte-specific secretory protein adiponectin critically depends on thiol-mediated protein retention. Molecular and cellular biology, 27(10), 3716-31

Wearsch, P.a. & Cresswell, P., 2008. The quality control of MHC class I peptide loading. Current opinion in cell biology, 20(6), 624-31.

Wei J, Gaut JR, Hendershot LM., 1995. In vitro dissociation of BiP-peptide complexes requires a conformational change in BiP after ATP binding but does not require ATP hydrolysis.J Biol Chem. Nov 3;270(44):26677-82.

Wetterau JR, Combs KA, McLean LR, Spinner SN, Aggerbeck LP., 1991. Protein disulfide isomerase appears necessary to maintain the catalytically active structure of the microsomal triglyceride transfer protein. Biochemistry. Oct 8;30(40):9728-35.

Wolin SL,.1994. From the elephant to E. coli: SRP-dependent protein targeting. Cell. Jun 17;77(6):787-90.

Yang, Z.R. et al., 2005. RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics (Oxford, England), 21(16), 3369-76. Yoshida, H. et al., 2003. A time-dependent phase shift in the mammalian unfolded protein response. Developmental cell, 4(2), 265-71.

Yoshida H, Haze K, Yanagi H, Yura T, Mori K. 1998.Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins. Involvement of basic leucine zipper transcription factors. J Biol Chem. Dec 11;273(50):33741-9.

Yoshida, H. et al., 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell, 107(7), 881-91.

Zhang, B. et al., 2003. Bleeding due to disruption of a cargo-specific ER-to- Golgi transport complex. Nature Genetics, 34(2), 220-225.

Acknowledgements

... e infine eccoci ai ringraziamenti.

Vorrei ringraziare il mio supervisore e director of studies Massimo Degano per la paziente supervisione durante tutto il dottorato. Come Virgilio con Dante, Massimo mi ha introdotto nella "selva oscura" della biologia strutturale incoraggiandomi a condurre il progetto in modo autonomo. Negli anni sono cresciuto sia come uomo sia come ricercatore imparando a gestire i piccoli e grandi problemi di NOI scienziati.
Ringrazio affettuosamente Roberto "Bob" Sitia per gli inesauribili consigli e per le numerose opportunità che mi ha concesso in questi anni.
Un ringraziamento speciale ai "cristallografi": Gianpiero, Elena, Paola, Claudia, Emanuela, Francesca, Benedetta, Arianna e Beatrice per avermi sopportato e supportato durante la mia permanenza nel laboratorio X-tal. In particolare a Gianpiero per le interminabili ed emozionanti ore al sincrotrone (e al loco mosquito) di Grenoble e a Paola per quello che sappiamo solo io e lei (inoltre riusciva a rendere meno asettico tutto il clima in laboratorio...purtroppo anche ciò che doveva rimanere sterile.....).
Un abbraccio al "ERp44 people" per avermi adottato nel lab-Sitia e per le stimolantissime discussioni scientifiche di questi anni. Il vostro infaticabile aiuto mi ha permesso di arrivare a fino a questo punto!!!!!
Grazie a senior del gruppo, Tiziana che mi ha dato accesso ai segreti più reconditi di ERp44 e a Silvia per i preziosi consigli e per le buonissime pizze. Un caro abbraccio a Leda e Stefy che mi hanno aiutato moltissimo a prendere confidenza con il progetto! Riccardo, Milena, Jose (one-way ANOVA sarebbe stato inarrivabile per me), Francesca, Martina, Eva ed Elena.
Grazie a "Guruji" Claudio Fagioli per il sostegno lavorativo e morale dell'ultimo anno. Grazie a te il WB non ha più segreti per me!!
L'immancabile un ringraziamento a Palma e Ana per l'insuperabile assistenza...la burocrazia sta a me, come la Kriptonite a Superman.
Certamente ringrazio ERp44 per avermi accompagnato tutti questi anni nella buona e cattiva sorte.

Tutta la mia gratitudine a Stefano Biffo, thirty party monitor, per avermi guidato in tutti questi anni con consigli e pacche sulle spalle: credo che senza Stefano non sarei qua a scrivere questi ringraziamenti.

La mia riconoscenza va a Luca Pellegrini, external supervisor, per i pochi ma ottimi incontri di questi anni: sempre forieri di consigli e ottime idee.

Scientificamente credo di dover ringraziare CC. Wang (Pechino-Cina) e Kenji Inaba (Fukuoka-Giappone) per le fruttuose collaborazioni!!!!!

Un grandissimo GRAZIE ai DiBiters che in questi anno hanno condiviso momenti di lavoro e "ludici". Specialmente i membri del gruppo Casari-Rampoldi,Musco, Valtorta e Biffo su tutti!!!!!!
A proposito di DiBiters, un immenso GRAZIE va a miei compagni di PhD Eugenio, Tanguy, Barbara e Patrizia!

Last but not least , Viviana! Grazie per tutto!

Un piccolo spazio lo lascio anche al MITICO autobus 925 per i meravigliosi viaggi!La mensa HSR ("buonissima"), i blackout elettrici (soprattutto quando si lavorava sotto cappa sterile), i freezer -80 (quando "saltano" durante il week end), la stanza di cristallizzazione (..si muore di freddoooooo li dentro), la macchinetta del caffè (compagna sempre presente...nei momenti belli e in quelli brutti), i DiBiT retreats + annual PhD students workshops (con relativa SPA e dinner party), i "flames" sulla mailing list del DiBiT, i "litigi" con i colleghi per decidere se accedere o meno l'aria condizionata, il GRANDISSIMO parcheggio del DiMeR (il mio meccanico e gommista ringrazia calorosamente), gli "algoritmi decisionali" per stilare la lista dei lab. meetings, i "think tanks" fra PhD students per migliorare le cose (aimè con pochissimi risultati pratici), il "micoplasma" che contamina le cellule in coltura(ma alla fine si è capito se fa male alle cellule?), gli anticorpi che non funzionano (maledetti HA e JAD1), le riunioni che iniziavano alle 19, i potentissimi "guardiani" (o addetti alle chiavi), l' Angelo, gli stressanti reports + PhD seminars e tanto altro ancora. ....

sarete SEMPRE NEL MIO CUORE

Vorrei ora concludere augurando in bocca al lupo:

- Ai PhD students (PhD tesi): cercate di divertirvi con il vostro progetto di dottorato (che alla fine durerà quasi 5 anni). E' un'occasione unica!!Ascoltate attentamente i consigli dei vostri supervisors, ma non lasciate che siano loro a fare la vostra ricerca: il progetto è vostro!!!!!!! Per quasi 5 anni mi sono divertito come un matto a pensare e fare esperimenti per capire

come funzionava **ERp44** (weekends e vacanze estive-invernali spesso "rimandate" in favore della scienza...). Vorrei continuare a lavorarci su ancora per tanto tempo!!!!!!!!!!!!!!!!

- Ai PhD students (social life): non vi chiudete in lab. a fare esperimenti. Andate ai seminari e interagite con più persone possibili!!!! Conoscere persone da tutto il mondo arricchisce e permette di confrontarsi e fare una ricerca migliore. E poi ci si diverte sempre la sera a bere una birra con i proprio colleghi. In questi anni ho incontrato molte persone al Dibit:bizzarre, intelligenti, particolari, strane, simpatiche, mediocri, false ma tutte mi hanno lasciato un ricordo!!
- A colui\e che seguirà il progetto su ERp44:ora tocca a te.
- Alla MERITOCRAZIA nell' università italiana (per quello che conosco, of course).....coraggio c'è la possiamo fare!!!!

Tralascio i ringraziamenti riguardanti argomenti non lavorativi: la tesi è già stata dedicata!!!!!!!!Posso aggiungere solo i miei gattini persiani (Izzi e Pandorino).Comunque ci sono state persone che, nonostante tutto, ricorderò per sempre in quanto hanno segnato indelebilmente questi anni, anche più della mia cara ERp44. Non credo ne siano consapevoli!

Stefano 27/07/2010